The Role of Emotionality in the Organisation of English Abstract Words on Second Language Lexicon

Yasir Almukhaizeem

PhD

University of York
Language and Linguistic Science

March 2021

Abstract

Abstract words (AWs) or imageless words refer to words which addresses something that is not tangible. Even though these words do not have physical existence, they can be categorised semantically into: emotional and non-emotional words. Whilst classical models offered valuable insights about AWs, the patterns of some AWs require further scrutiny. More studies have proven the role of emotionality in the representation of these words among monolinguals (Kousta *et al*., 2009; Vigliocco *et al*., 2010), but few studies have explored the role of emotionality on L2 speakers. This thesis examines how these words are rated and stored among L2 speakers of English ($n=60$) with different levels of language proficiency compared to monolinguals ($n=30$). Particular attention is given to examine the effect of emotionality in storing these words. Data obtained suggest that not all AWs are scaled and stored in the same way, as some AWs showed different patterns of emotionality intensity as well as retaining these words in the memory. The result showed differences at the word level rather than in the type of the lexicon. Emotional words were rated higher than non-emotional words across groups. Non-emotional words displayed different patterns in terms of emotionality. As for storing these words at the retention stage, more patterns emerged between emotional and non-emotional words. The thesis concludes that the representation of AWs is affected by the semantic characteristics of these words, the design of memory tasks and the degree of word difficulty. Emotional intensity may enhance the recallability of emotional words, but the nature of memory tasks and the type of AWs affect this relation. However, the level of language proficiency has proven to be less effective among L2 speakers in establishing emotionality. The establishment of emotionality and memory retrieval are two independent processes which are unaffected by the speakers' linguistic abilities.

Table of Contents

List of Tables

List of Figures

Dedication

I dedicate this work to my beloved family

My mother and late father

Sarah and Saad

Acknowledgements

First of all, I would like to express my thanks and gratitude to Allah -Almighty - for granting me determination, patience and time to do this work. Without His strength, it would not have been possible to finish this work on time. Sincere thanks and appreciations go to my supervisor Dr. Eytan Zweig for his help in supervising me for four years. Dr. Zweig was an ideal supervisor with his time, and care for his students from the date of commencing until the viva. During my PhD journey, he was keen to provide me with constructive feedback, endless encouragement and valuable guidance. Warm appreciation goes to my thesis advisory panel (TAP) member. Prof. Tamar Keren-Portnoy for her feedback and guidance. Undeniably, she inspired me with valuable recommendations and motivating ideas for my project in TAP meetings as well as through PhD seminars. Both of them helped me develop a small proposal to a PhD dissertation in linguistics, and introduced me to the field.

There are people in mind for whom I have been truly thankful throughout my PhD life. At the top of my list, Prof. Marilyn Vihman was very supportive at early stages of my thesis. Prof. Vihman was the chair of my progression panel for two years. Truly, her ideas solidified the methodology chapter and enriched the structure of the thesis overall. My sincere appreciation is extended to Dr. Heather Marsden for her recommendations. I should not also forget Prof. Jim Milton for his comments. Further thanks go to the postgraduate administrators, Ms. Gem Kirk, Mr. Geoffrey Krause and Mrs. Jane Crowther for their constant help in making my job easy, starting from academic admission, visa registration, data collection and finally thesis submission. Also, I need to raise my appreciation for the library staff at the University of York for helping me finding resources whenever I need them. They showed support with encouraging words, especially when they saw me working late at night in the JB Morrell Library. Huge thanks go to Mr. Tony Wilson and Mr. Sam Hazeldine for their effort and time in IT support team as well as locating theses and papers in the university repository and databases. Both of them helped me a lot during the lockdown of the Coronavirus in order to meet deadlines.

I am truly grateful to King Saud University for granting me a scholarship to do a PhD in linguistics. I am thankful to Dr. Salem Albuhari, Prof. Saad Alhashsh, Prof. Mubarak Alkhatnai, Prof. Abdurhman Almansour, Prof. Nasser Almansour, Dr. Ghazi Alotaibi, Prof. Mufleh Alqahtani, Dr. Saleh Alqahtani, Prof. Ibraheem Alqarni, Dr. Sami Alsulami, Dr. Abdullah Althunayan and the current dean Prof. Sultan Alharbi. Special thanks for Mr. Abdurhman Alghamdi and Mr. Abdulkareem Alotaibi from Human Resources Department at King Saud University. Further thanks go to the Saudi Arabian Cultural Bureau in London for their assistance during my stay in UK. Prof. Amal Fatani has been very supportive, especially in the second wave of the pandemic. I should not forget to thank staff of King Salman Central Library in Riyadh for their support during my visits to Saudi Arabia.

Warm wholehearted appreciation is raised for my family in Saudi Arabia. I cannot truly describe my emotion in expressing my gratitude to my mother and late father who passed away before my PhD. I will be forever indebted to their sacrifices in effort and time. They always believe in me even when I have doubts and fears. I ask Allah to reward both of them with blessings. I will be forever grateful to my mother's night prayers and strong wishes, asking Allah to make my PhD journey easy even when she was hospitalised for more than two months in three different hospitals in Riyadh. I am speculating after the viva, she will ask me about the due date of coming back to Saudi Arabia only! True thanks and appreciation also go to my brothers and sisters: Amal, Aljawher, Abdullah, Sami and Ibtesam. During my stay in the UK, they have been very supportive, filling my temporary absence in the family. I will be forever thankful to every one of them.

This is not the end of my list. I am grateful to my beloved wife who assisted me while I was so involved in my studies. Undeniably, she expressed no frustration whatsoever while we were away from Saudi Arabia, especially during the pandemic. Although those days were tough and discouraging due to a total self-isolation and lockdown in Britain, she created an encouraging and lovely environment to study in York. I am also thankful to my daughter Sarah.

I will always remember her saying “Daddy! What time will you come back home? I need a chocolate bar!”. A special thank goes to my new baby born Abdulaziz who was born five months before my viva. His smiley face encourages me to keep working on the thesis until five in the morning. Indeed, I am really grateful to all of them and I know they have made enormous sacrifice.

This acknowledgment is not finished without sending warm appreciation to my friends in York. I am truly indebted and thankful to Khaled Alasmary, Suilam Alataouy, Anwar Aljadani, Jassas Alotaibi, Ibraheem Alwafy, Eleanor McCullough, Almoutazbellah Hassan and Zuhair Qabsi. Special thanks are extended to Ali Thafer and his son Suliman for their help from the moment I came to York. They made great effort to find me the right environment to study. Thanks also go to Ms. Deana Overton from Quantum Real Estate office in York for her assistance in finding me the right place to live for more than four years. I am sure I will not forget 6 Heron House (YO193DE) in York for the rest of my life. Indeed, it was a wonderful neighbourhood to live.

Last not the least, I am truly thankful to my English and Saudi participants in Brighton, Exeter, Leeds, London, Newcastle, Sheffield, Southampton, Swansea and York. I am thankful to their contributions, suggestions, and comments which made my study stronger. Huge thanks for my friends, Ali Alayany, Abdullah Alharby, Mohammad Alshjajer, and Wa’al

Alqurashi, Mohammad Bani, Abdullah Ba Sunboul and Abdullah Iskandarani for their effort in organising my data collection trips to these cities to recruit more participants. Without their tremendous support, this thesis would have been hard to accomplish on time.

Declaration of Authorship

I declare that this thesis is a presentation of original work and I am the sole author. It has not been previously accepted or awarded any academic degree from the University of York or elsewhere, either partially or wholly. All resources and quotations have been referenced and acknowledged at the end of the thesis as References.

I also declare that part of my thesis was presented at the following conferences (listed from the most recent to the oldest ones):

- 31 European Second Language Acquisition conference, held at the University of Fribourg, Switzerland on 26th August 2022;
- Manchester Forum in Linguistics, held at the University of Manchester, United Kingdome on 28th April 2022; and
- Linguistics and English Language Postgraduate Conference, held at the University of Edinburgh on 3rd May 2019.

Chapter 1

Introduction

1.1 Introduction

This thesis studies the representation of *abstract words* (AWs), words which refer intangible ideas rather than physical existence. The nature of AWs reflects the instability of memory patterns among monolinguals compared to words which can be observed by senses, like *concrete words* (CWs) (Altarriba *et al.*, 1999). It is suggested that the degree of *emotionality* (emotional experiences, affective status, and past memories) affects the recallability of these words (Kousta *et al.,* 2009, 2011; Vigliocco *et al.*, 2010, 2013). Proving the impact of emotionality among *second language* (L2) learners would indicate that this factor plays a significant role in remembering various types of AWs among L2 speakers. Furthermore, these patterns are presumed to be complicated among L2 learners. The current research intended to investigate this assumption further and to determine its accuracy. Therefore, it aimed to explore the effect of emotionality on the recallability of various types of AWs between monolinguals and L2 speakers. The current chapter provides an overview of the thesis and the research, introducing the research background, statement of the problem, the significance and the context of the study. Further details will be provided on the research questions and hypotheses, alongside the conceptual direction of the study. Key concepts and terminology related to AWs are then explored before outlining the structure of the thesis.

1.2 Overview of the Study

1.2.1 Research background

AWs have been the subject of debate for over fifty years, with attention being paid to the memory patterns of these words when compared with other types of words since the early 1970s (Paivio, 1971a). At the beginning of the 1980s and 1990s, semanticists and psychologists disagreed over the behaviours of AWs, criticising the existing literature on how these words were modelled (Altarriba *et al.*, 1999; Johnson-Laird, 1983). Later studies noted that it is hard to assume that these words can have the same features in the way they are understood because they have no physical and tangible referent (Barsalou & Wiemer-

Hastings, 2005; Guthrie, 2012; Khokhlova, 2014; Kousta *et al.*, 2009). Based on this stance, various studies tried to identify what abstract means in this context. From a semantic perspective, AWs refer to something that does not have any physical existence (Johnson-Laird, 1983; Khokhlova, 2014). The degree of abstraction varies in AWs, with some AWs refer to affective states and previous experiences since AWs denote intangible concepts. In addition to their semantic meaning, some AWs have different levels of emotional intensity (Kousta *et al*., 2011; Vigliocco *et al.*, 2013). To clarify, an AW with a high emotional arousal, like *love,* is perceived with diverse emotions, like a mother hugging her child or a married couple conversing (Farley *et al.*, 2012). Other AWs, like *knowledge,* have no or limited emotions representing its meanings. Alternatively, some AWs can be linked to images instead. These latter types of AWs reflect different processing patterns in the lexicon, whether this is based on linguistic knowledge alone or other factors that affect the representation of these words. Taken together, words, like AWs, are structured in the lexicon according to two types of knowledge: linguistic and affective (Della Rosa *et al.*, 2018). By looking at the examples of AWs stated earlier, we can see various emotional patterns reflecting how the lexicon deals with types of AWs. Even the patterns of the words, *love* and *knowledge*, in terms of *imageability* (images shared with a word, either static or dynamic) and *concreteness* (a level to which a particular word denotes a tangible object), vary without any consistency (Altarriba *et al.*, 1999).

AWs have been studied from a range of different perspectives, including psychological, neurolinguistic and psycholinguistic. From a psychological perspective, there has been a tendency to assume that the lexicon deals with AWs differently and, thus, AWs show varied behaviours in terms of different word norms. Altarribia *et al*. (1999), for example, investigated three norms in AWs: imageability, *availability of context* (information from the speaker's knowledge to explain the meaning of a word) and concreteness. Their study considered emotional words a sub-category of AWs. The outcomes of the three norms proved the variations between AWs. One of these outcomes, for instance, was the variation between emotional and non-emotional words in terms of imageability in the sense that emotional words had a higher level of imageability than non-emotional words. From this, words with a high level of imageability, like emotional words, remembered better. A neurolinguistic perspective provides another approach of studying AWs. Studying AWs from this perspective has underlined various neurological activations of AWs, through using neuroimaging, neurophysiology and magnetoencephalography (Pulvermüller, 2013)[1]. By focusing on the biological side of processing AWs, neurolinguistic studies have observed more brain activation for emotional

[1] These are three ways to measure the activation in the brain and the nervous system but from a different perspective. Neuroimaging refers to using different techniques, like *magnetic resonance imaging* (MRI) or *computerised tomography* (CT) scan to examine or diagnose the brain if there is any abnormal brain activation. As for neurophysiology, it studies the nervous system of the body to assess any nervous disorders. Magnetoencephalography focuses on measuring the electric current of the brain by observing the magnetic fields.

words than non-emotional words (Pulvermüller, 2013). This finding indicates that both emotional and non-emotional words are represented the same in the brain. In psycholinguistics, the cognitive processes involved in perceiving and recognising AWs is another dimension of studying AWs. This perspective aims to show how AWs are accessed, recognised, stored and understood (Ferré *et al.,* 2010, 2013, 2018). AWs are considered to reflect a different mechanism, especially when examined as part of a bilingual lexicon. The characterisation of the lexicon as monolingual or bilingual influences the way AWs are represented in the lexicon. Language proficiency and how AWs are stored in different memory templates, among other factors, can also affect how they are organised in the lexicon. There have, therefore, been a range of perspectives which have utilised a range of approaches and testing conditions to facilitate better understanding of AWs. These perspectives provide diverse insights into AWs and demonstrate the importance of understanding these words, and how they are processed by different people in different contexts. Examining how these words retained in the memory, and whether there is an effect of remembering with high level of emotionality is still missing.

1.2.2 Statement of the problem

Existing literature indicates that emotionality as a psychological factor is involved in the recallability of AWs. It has been shown, for instance, that some AWs with high levels of emotionality behave, like CWs, in their recallability

(Altarriba & Basnight-Brown, 2011; Altarriba & Bauer, 2004; Altarriba *et al*., 1999; Kousta *et al.,* 2009, 2011; Vigliocco *et al.*, 2010, 2013). In fact, the impact of this factor plays a role in retaining AWs in the memory since memory is the place for storing words. Without it, some AWs would most likely be poorly remembered. The evidence for such claims, however, has tended to draw on the memory patterns of monolingual speakers, therefore, studies on storing different types of AWs in the memory among L2 speakers are still inadequate to examine factors affecting memory retention (Ferré *et al.,* 2013). Furthermore, there is a lack of consensus surrounding the role of emotionality within AWs. Of the limited research in this area, it has been shown that words acquired with emotional associations among early L2 speakers (speakers who acquired their L2 earlier in their lives compared to late L2 speakers) are remembered better than late L2 speakers (Ferré *et al.,* 2013; Pavlenko, 2005). Additionally, research by Anooshian and Hertel (1994) and Ayçiçeği and Harris (2004) recognised the role of emotional intensity for both L1 in L2 speakers. It was not clear, however, whether L2 speakers observed the differences between AWs based on the strength of emotionality in the lexicon of the two languages. This factor affects the way these words are retained in the memory. The emotional knowledge among L2 speakers is still developing, as the effect of this factor varies when the linguistic abilities of L2 speakers differ. Therefore, it is said that the organisation of the lexicon differs between monolinguals and L2 speakers (Milton, 2009). Whilst there is some research on L2 speakers, the current literature does not sufficiently investigate the level of emotionality of AWs among L2 speakers, and has not compared the outcomes with monolingual speakers in a single study using one language. Moreover, the behaviours of different types of AWs among L2 speakers at different memory stages are still unclear. Finally, the characteristics of AWs make studying these words challenging, and further consideration of the different types would be beneficial for developing a clear understanding to these words (Wiemer-Hastings *et al.,* 2001).

1.2.3 Significance of the study

Conducting further investigation into AWs was considered vital for three reasons. First, previous research has concentrated on limited memory types and most importantly treating all AWs as a homogenous group (Paivio, 1971a, 1973). Relying on the results from these testing environments for examining AWs is arguably insufficient. Memory involves more than just remembering words. It underlies a sequence of stages: encoding, retaining and retrieving words from memory. Even recent memory studies of AWs have identified theoretical and methodological concerns surrounding evidence in existing literature. At the theoretical level, for instance, AWs in the literature have been treated equally, with claims that they are represented in the lexicon similarly (Paivio, 1971a, 1973). This view emerged in the late 1970s in response to the need to study AWs since studies at that time the patterns of AWs had not been sufficiently studied, requiring more work (Gorman, 1961). Although studies on these words have gradually increased, research has continued to impose some constraints on the type of lexicon and the selection of AWs. Furthermore, whilst there have been some important developments in the field, especially in the work of Vigliocco and Kousta (2009, 2010, 2011, 2013) who identified the embodied emotionality model of AWs, these two constraints still need to be addressed in terms of their effect on the lexicon.

Second, it was felt that further research was required to examine the recallability of various AWs among L2 speakers and determine whether monolinguals and L2 speakers have similar/different memory patterns. AWs remain challenging, not just in one language, but in a lexicon with two languages. These challenges rely on the fact that AWs do not have fixed features, as some have different levels of images and concreteness (Altarriba & Bauer, 2004; Altarriba *et al.,* 1999). Semantically, some abstract emotional words may not have one word as a translation across languages (Dewaele & Pavlenko, 2002). A word like, *Zeitgeist* in German, refers to ideas, faiths and intellects of age. The behaviours of these words in the memory solidify the complexity of these words, as it is hard to generalise what we already know about AWs from the literature (Altarriba & Bauer, 2004; Altarriba *et al.,* 1999). L2 learners, in particular, struggle to learn various types of AWs, regarding translation and memory, and these words remain passive in their lexicon. The indications in Altarriba's work (1999) raises questions around how AWs are represented in the lexicon, and whether we can model the recallability of these words among L2 speakers. Therefore, studying AWs is crucial for explaining the mechanism for storing AWs in memory among L2 speakers, where two languages are stored together in one place.

Third, it would facilitate our understanding of introducing these words to L2 learners. Being a lecturer of English helps in observing how L2 speakers remember AWs. During the learning process of AWs to L2 speakers, it has become evident that some types of AWs are remembered better than others with no obvious explanation. Thus, developing knowledge here, through unravelling the complexities of AWs and comparing understandings of AWs between monolingual and L2 speakers, may facilitate teaching and learning practice in languages. Realising the importance of these three reasons for studying AWs, conducting a study about these words becomes significant.

1.2.4 Context of the study

The current research examines the impact of emotionality on the recallability of AWs, and whether the level of emotionality within AWs is similar among L2 and monolingual speakers. As touched on above, there is a lack of consensus little attention is paid in the literature to the strength of emotionality among L2 speakers, indicating assumed equal levels of emotionality between L2 and *first language* (L1) (Ferré *et al.,* 2013). At the word level, there were no ready-made corpora to measure the emotionality of all AWs among L2 speakers, especially if the selection of words considers various variables for emotional and non-emotional words. Choosing AWs according to particular criteria can provide a helpful environment to measure emotional intensity. Therefore, modelling the emotional intensity between various types of AWs is essential for observing the difference between these words. On the other hand, there have been some limitations in generalising this level of emotionality to all L2 speakers in all memory types. Consequently, the study aims to examine the recallability of various types of AWs in two memory conditions and observing the emotional intensity linked with these words. This approach is valuable for studying the structure of the lexicon and characterising how memory is processed (Rogers *et al.,* 2018).

The current project contributes to the existing work by examining two samples of L2 speakers to investigate the use of emotionality in AWs, and how it impacts their ability to retain them. It is, therefore, difficult to draw any conclusions by only studying advanced L2 speakers. Having identified this gap in the literature at the methodological level, the study examined both beginners advanced L2 speakers to analyse whether or not there is a difference between the retention of AWs based on the words' level of emotionality. This is to offer explanations for the instability of memory patterns in the existing literature. As such, the current study aimed to investigate the durability of emotional and non-emotional words in the memory by examining whether the establishment of emotionality helps the speaker to retain these words temporarily or permanently. With the various levels of emotionality in AWs stated in the literature (see Chapters Two and Three for details), different AWs were selected for this study in two memory conditions designed to examine the behaviours of these words in single and bilingual language environments. Since the role of memory is to retain words, it is not yet clear whether emotional words are remembered better than non-emotional words at the retention stage (whether STM or LTM). It is hard to generalise the memory patterns in the literature to all memory types. The two L2 groups represent the experimental group, and the monolingual group represents the control group. The outcomes of the three groups will be examined.

1.2.5 Research hypotheses, questions, and aims

The study's hypotheses were based on the researcher's observations and extensive reading in the literature (see Chapters Two and Three for further details). The role of emotionality has proven to be vital cognitively among monolinguals in processing AWs. There are a few indicators in the literature that also pinpoint the role of emotionality among L2 speakers (Altarriba & Bauer, 2004; Altarriba *et al.*, 1999; Ayçiçeği & Harris, 2004; Siakaluk *et al.*, 2014, 2016). However, bearing in mind their level of language proficiency, the representation of AWs among L2 speakers is still lacking. Consequently, this thesis argues that, with the influence of both emotionality and language proficiency, the memory behaviours of AWs will be different between the two levels of L2 speakers. The outcomes of this study will attempt to prove these suppositions. Thus, the study hypothesises that emotionality plays a role in the representation of AWs among L2 speakers compared to monolingual speakers, distinguishing between both various types of AWs and L2 speakers. At the same time, this hypothesis is non-directional in not deciding on the distribution of emotionality and language proficiency on all memory patterns of AWs. Therefore, this hypothesis assumes that:

1. The distribution of emotionality varies at the word level between emotional and non-emotional words on the one hand, and between L2 speakers and monolinguals on the other; and

2. Emotional words are hypothesised to be better remembered compared to non-emotional words in both STM and LTM among L2 speakers.

The research aimed to address the following research questions:

1. How do L2 speakers of English, whose linguistic abilities vary, rate emotional and non-emotional words in terms of their emotionality compared to monolinguals?
2. What are the underlying memory patterns of emotional and non-emotional words in STM and LTM as L2 speakers have various linguistic proficiency?

3. Is there a relationship between the level of emotional intensity linked to emotional and non-emotional words and the recallability of these words? What are the factors that could affect this relation?

The aims of the current study were:

1. To examine the distribution of emotionality of emotional and non-emotional words, and between positive and negative emotional words, among monolinguals and L2 speakers;

2. To study the memory patterns of emotional and non-emotional words and whether or not they are similar in STM and LTM; and

3. To link the outcomes of the patterns of emotionality to the recallability of emotional and non-emotional words to determine the impact of emotionality on the recallability of these words among monolingual and L2 speakers.

1.2.6 Conceptual direction of the study

The conceptualisation of the emotionality model by Kousta and Vigliocco (2009, 2010, 2011, 2013) is central to this study. This section explores the background to its developments, core components, potential limitations and how the emotionality model was applied within the current study. Thus, it outlines the conceptualisation of AWs. As mentioned previously, for instance, the representation of AWs was first addressed by psychologists, like Paivio in the 1970s, and it is helpful to begin here by outlining the classical models of

AWs in this field since they form the background to the development of the model[2].

Historically, the interest in studying AWs dates back to Aristotle's *De Anima* (350 BC) and Plato's *Republic* (375 BC). They looked at the intangible notions of the unseen world (Guthrie, 2012). Over time, this interest in AWs grew to encompass their semantic features, and how they are stored in the memory. Paivio's DCM (1971a, 1973, 1991) and Schwanenflugel and Shoben's context-availability model (1983) addressed the mechanism of storing AWs in the memory. These models sought to characterise the patterns of storing AWs compared to CWs through a series of memory, scale rating of imageability and availability of context tasks. Concerning the variations of AWs, there has been little to say about the representation of different types of these words in these models. These models suggest that the representation of words in the memory can vary in nature between visual (Paivio, 1971a, 1973, 1991) and verbal (Schwanenflugel & Shoben, 1983). Thus, according to these studies, AWs are represented only at the verbal linguistic level with no or limited visual images shared with these words. Importantly, these models compared the patterns of AWs to CWs, and did not provide a reason, as to why

[2] The reference to Paivio and colleagues' work in discussing the representation of AWs is limited in the thesis to two main points. The first point is the influence of their work to reconsider the representation of AWs since they did not provide a sufficient explanation about the instability in remembering these words. Therefore, the reference here is to the historical development of studying AWs, starting from Paivio and colleagues and finishing with Kousta and Vigliocco. Secondly, there will be another reference to the dual representation of CWs later in the thesis, as the current study used CWs in three experiments as a control group.

AWs were so poorly remembered (Paivio & Csapo*,* 1969; Schwanenflugel & Shoben, 1983; Schwanenflugel & Stowe, 1989). Taken together, these studies suggested that the absence of the visual level could be the reason for this memory pattern of AWs. They assumed that the visual and verbal levels are interconnected, as the visual representation focuses on processing the visual non-linguistic elements of words, unlike the verbal representation, where the focus is on the linguistic elements. The role of the lexicon then focuses on the interplay between the two levels, depending on the types of words. A word, like *chair*, is more likely to be represented at both the visual and the verbal levels. As a result, CWs are processed faster than AWs because AWs tend to be represented only at the verbal level. Therefore, the mechanism of storing AWs in the memory lacks a secondary level of representation which is why AWs are more difficult to remember than words with tangible reference. At the contextual level, the structure of DCM was made in a primarily monolingual environment where speakers had diverse levels of imageability and availability of context (Figure 1)[3]. Based on DCM, understanding the representation of these words is limited to the role of imageability and context availability. The

[3] The squares in the verbal system, according to DCM, are connected because a particular word needs more than letters to be stored. Circles in the visual system, on the other hand, represent an image schema for a particular word. Therefore, these circles are connected since DCM suggests that an image represents each word at the visual level. The design of this figure was made by the researcher which was based on DCM discussed in (1971a).

role of other factors, like the emotional intensity of words, was not discussed in these models.

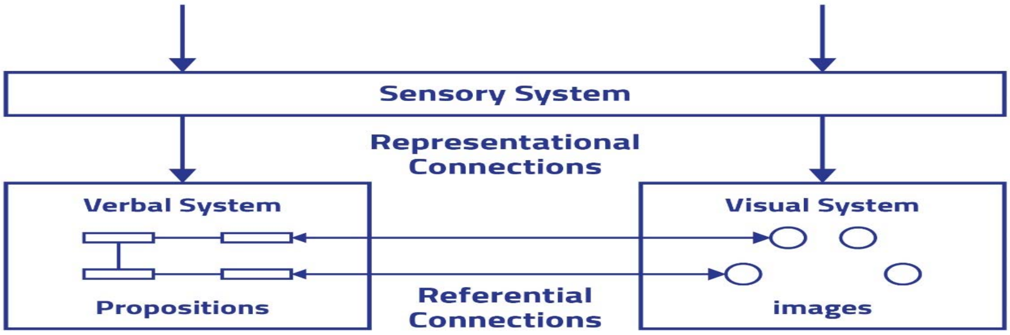

Figure 1. *Structure of Paivio and colleagues' DCM (Source: Paivio et al., 1971a).*

Questions about whether or not a bilingual speaker has two encoding processes have been raised in the literature. This issue was addressed by Paivio and Lambert (1981), who questioned whether bilinguals have two separate memory stores. What can be observed is that the results of memory, translation and picture-naming tasks support the conceptual frame of DCM, not only among monolinguals but bilinguals as well. To be more specific, the semantic characteristics of the word are also vital because memory differentiates between words with(out) similar origin or what are termed as cognate and non-cognate words (e.g., *a table* in English *and la table* in French are cognates). The process of remembering cognate words, either AWs or CWs, tends to be easier and faster compared to non-cognate words (De Groot, 1992). There are shared nets and nodes in the brain between words which tend to be activated in two different memories compared to non-cognate words. Paivio and Lambert

(1981) concluded that DCM had another version for bilingual speakers. Depending on the type of words, DCM for each language can work unconnectedly or jointly. CWs are an excellent example to clarify the link between dual representation and the type of words. A word, like *book,* has imagery that has almost similar descriptions between languages. Either in English or any other language, an object with pages with hard covers is termed *book*. At the verbal level, it is termed *book* in English, but in Arabic, it is termed /kitāb /كتاب. The DCM for a bilingual speaker of English shares the visual representation of the word book, but the verbal representation is different. This difference is attributed to the difference between English and Arabic in pronouncing and writing the word *book*. For a word, like *computer*, which has the same imagery and pronunciation in English and Arabic, the representation of this word is shared between these two languages (Figure 2). At this point, we can assume that DCM is shared between the two languages, and therefore, the recallability of this word is better compared to non-cognate words. For AWs, the representation of these words happens at the verbal linguistic level which is connected to their meanings, with low or no shared imagery. The low level of imagery, according to DCM, means that AWs are represented at one level (Clark & Paivio, 1991; Paivio & Desrochers, 1980).

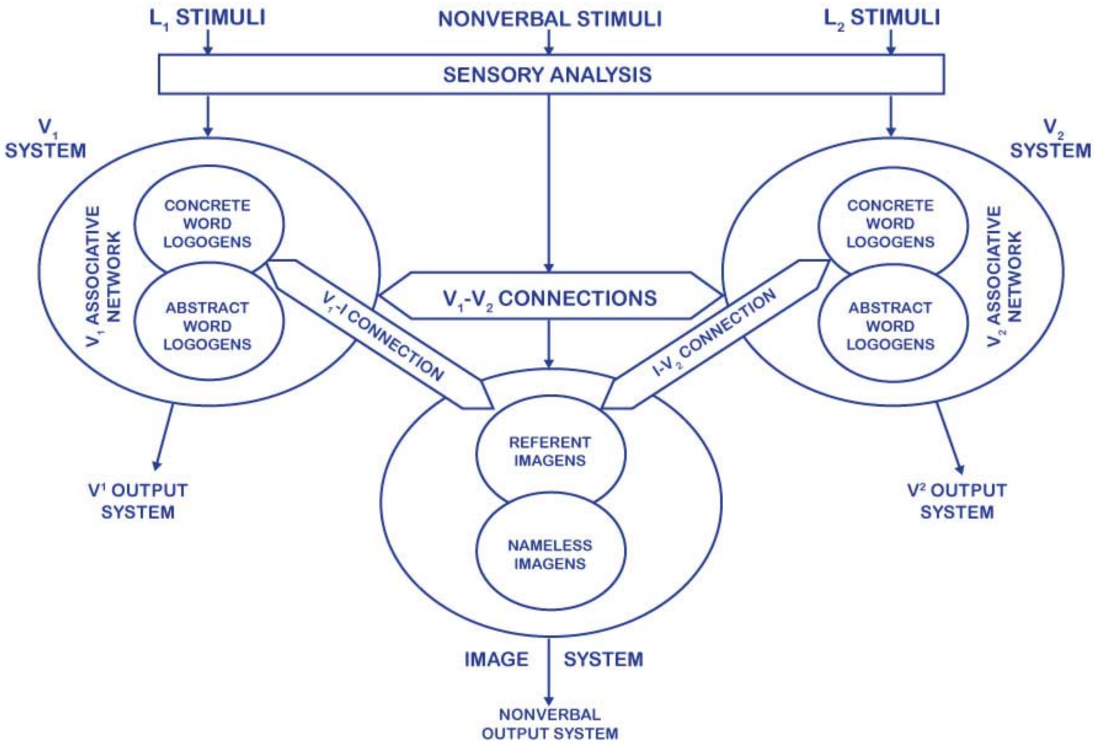

Figure 2. *Bilingual version of DCM (Source: Paivio & Desrochers, 1980).*

Among L2 speakers, AWs are also verbal, and this level of representation signifies the meaning of the word only. DCM sees that AWs have less or no visual associations at all, and therefore they are poorly remembered compared to CWs. The singularity of representation happens since no visual associations could facilitate the meaning of AWs. According to the definition of AWs in DCM (Paivio, 1971a), AWs do not have any physical existence in real life, and no mediator can represent the meaning. At the linguistic level, there are linguistic differences between L1 and L2 in terms of encoding these words at the phonological level. A word, like *idea,* as an AW, does not have a physical existence in real life. For an L2 speaker of English whose L1 is Arabic, the word *idea* refers to فكرة /fikra/ which is only encoded at the verbal level, unlike the word *computer*. In Figure (2)[4], AWs in both DCM are represented at the verbal level (logogens). However, when an AW is a cognate between the two languages, encoding this type of word is similar, except for minor phonological differences. An example here is the word *noble* in English, نبيل/ nabi:l/ in Arabic.

Like any model, there were concerns over the validity of this mechanism of representing words in the memory. These concerns can be classified into two main parts. The first part relates to the mechanism of storing words as verbal or nonverbal representations. Coordination between these levels remained blurred in the model, as it is hard to assume that no changes or transformations happen to these two levels with the frequent use of a word over time. Many factors could affect this representation, like the type of memory, and whether words are processed in a monolingual, bilingual or multilingual environment. It still needs to be determined whether or not the duality of representation remains the same in these different parts of the lexicon for the same word, mainly used by multilingual speakers. Moreover, one of the factors that influenced the encoding stage is the age of speakers and their ability to create mental imagery. This encoding process is assumed to be progressive from early

[4] By looking at this figure, we can see the representational system of L2 speakers is shared in labelling words with tangible references at the visual level by referring to similar referent imagens in L1. For AWs, there is no connection between L1 and L2 at the visual representation. Each language has a separate verbal representation (logogens).

childhood until adulthood. DCM did not provide more details about the characteristics of this change or how the visual representation may change (Paivio, 1991). Also, there was no change in the model to address this issue, considering age as a factor which can affect the creation of imagery among either children or adults.

The second set of concerns relate to the variation of AWs. Since the 1970s, for instance, research has argued against the conceptual frame of AWs in Paivio's model (Johnson-Laird, 1983). The variation in AWs was actually recognised prior to the introduction of the DCM. We can argue that assuming that all AWs behave in the same way in memory is illogical, as these words describe something that is not observable by human senses and has various levels of abstraction. The difference between emotional and non-emotional words is an example here. DCM did not, therefore, provide a clear discussion about AWs. The role of imageability in DCM can help in finding the differences between AWs and CWs.

Crucially, DCM did not tackle the difference between the form and content of a word. The model was unclear about whether the verbal representation refers to abstract nouns, adjectives, verbs or adverbs. At the same time, there was no discussion of CWs reflecting abstract meaning, and whether they are similar to AWs in terms of their levels of representation (e.g., *cancer*). This led to imprecise results about AWs with no further attention paid to form and content. This was because DCM was designed primarily, as a framework for psychology, and had less emphasis on linguistic variations in the selection of words. From a linguistic point of view, processing words, including AWs, varies between different word classes. By looking at the corpus used in DCM studies by Paivio in the 1970s, we can see that the AWs used in this model raised concerns because DCM mixed nouns and adjectives and then generalised the results to reflect all AWs. We can see that the variations between AWs were not fully addressed which necessitates further investigation. This is one of the concerns regarding the generalisation of DCM findings regarding AWs, which were not entirely accurate.

Kousta *et al*. (2011) and Vigliocco *et al.* (2013) addressed the representation of AWs in the lexicon from a different viewpoint. They were not entirely satisfied with Paivio's conclusions about AWs. They focused on a broader perspective beyond just a static image. That is to say that the speaker's knowledge about AWs is not limited to linguistic knowledge. The representation of these words goes beyond this type of knowledge, and increases over time through the acquisition of experiences, interactions and – most importantly – emotionality. This view combined the representation of AWs with the level of emotionalit which is one form of embodied knowledge about these words. Subsequently, it became recognised as the embodied emotional knowledge model (Kousta *et al.*, 2011; Vigliocco *et al.*, 2013).

Ultimately, this view linked the representation of AWs to our acquired experiences, previous memories and affective states. A question emerges about the characteristics of knowledge attached to AWs. This attracted Kousta and Vigliocco's attention (Kousta *et al.,* 2011). They believed that this knowledge includes sensory affective states, suggesting that the representation of AWs is based on the person's emotional arousal. They claimed that these words are processed internally by identifying their semantic meaning in the lexicon, and then externally by perceiving experiential knowledge, affective status, and old memories that symbolise their meanings.

The behaviour of sharing the meaning of words with imagery in the lexicon has been proven psychologically and neurolinguistically (Barsalou & Wiemer-Hastings, 2005; Martin & Ellis, 2012). The role of the lexicon does not stop at this stage. It involves an ongoing process of comprehending the meaning of AWs, focusing on semantic features and associations of these words. It can create a network of words with the same features. Therefore, after differentiating AWs semantically, the lexicon takes another step in classifying any information, experiences, memories and affective states related to AWs (Kousta *et al.,* 2009; Kousta *et al.,* 2011). These words then behave differently, depending on the acquired embodied knowledge, and can be categorised into two types of AWs. The first type is understood with more embodied knowledge obtained from situations, contexts and experiences. These words are linked by different associations and work as a tool to represent the meaning jointly. These words are termed emotional words. A second type of AWs, or non-emotional words, is constrained by a minimal understanding of their semantic meaning with less or no indication of storing them with embodied knowledge. These words are stored differently in the lexicon compared to the first type. In both types of AWs, identifying words begins by perceiving the word visually or aurally. The lexicon processes these words to encode any relevant associations, either linguistically or non-linguistically. However, the process becomes complicated when another action must be done promptly due to perceiving an AW. In a more precise example, *joy,* and *depression* are two different AWs that represent two extremes of positive and negative emotional intensity. Sometimes, some positive emotional words are perceived with a smile requiring a facial movement. On the contrary, words, like *concept* and *thought,* reflect no or limited emotion, and are therefore disembodied from any knowledge (Barsalou & Wiemer-Hastings, 2005).

The characteristics of non-linguistic knowledge vary in their nature, including emotions, experiences and memories. For example, when a word, like *passion,* is heard, the node is activated by any embodied knowledge which could include love or a romantic exchange between husband and wife. Data from Naccache *et al.* (2005) revealed that processing negative emotional words is done not only at the linguistic level, but also involves motor movements in the face and some brain activations in the amygdala. As such, some AWs have rich semantic associations, as they tend to create more links with words with the same features (Deyne & Storms, 2015). Overall, there are no fixed prototypes of an AW. Instead, prototypes vary within AWs and across languages. The features of this knowledge can be previous memories that represent the meaning in mind, including emotional intensity linked to AWs (either positive or negative) and affective states, which include emotional experiences, shared with AWs.

Overall, the representation of words in the lexicon is affected by many factors, such as imageability, context-availability and emotionality. All of these factors stated earlier can be tested using a rating scale. Toglia and Battig (1978) were the first to rate AWs according to imageability, highlighting the importance of this factor in the representation of words. Subsequently, Schwanenflugel and Shoben's model (1983) included a scale which rated AWs by their context availability. These scales have been used to verify the fact that AWs are influenced by these factors, and the data have provided a reason to claim the effect of imageability and context-availability on the representation of AWs in the lexicon. Characterising the distribution of the embodied knowledge within AWs also includes other tasks, like generative, memorisation and lexical decision tasks. For generative task, it has been used to examine the effect of emotionality on AWs. It is based on asking the participants to generate words or descriptions of AWs. It is also known as a feature listing (Barsalou & Wiemer-Hastings*,* 2005; Marmolejo-Ramos & Dunn, 2013) or word association task (Deyne & Storms, 2015). The results indicate that associations of AWs are linked more to emotional words and less to physical objects. According to Marmolejo-Ramos and Dunn (2013), AWs are shared with more emotional affective knowledge and previous experiences in this context. As for memory patterns, the findings of Altarriba and Bauer (2004) and Altarriba *et al.* (1999) clarified that emotional words were remembered better compared to non-emotional words. Their two studies hypothesised that emotional words behave, like CWs, in memory patterns due to their emotionality, compared to other types of AWs.

In the 2000s, Kousta and Vigliocco noted the lack of data on AWs and proposed this model for the representation of words, based on experiential context and emotional knowledge (Kousta *et al.,* 2009; Kousta *et al.,* 2011; Vigliocco *et al.*, 2013). Their studies supported the existence of embodied knowledge within the acquired words. The model, in fact, is novel in considering the role of emotionality in organising AWs in the lexicon (Siakaluk *et al.,* 2014). To be more specific, different AWs ranged from high to low emotional intensity. According to Siakaluk *et al*. (2016), this factor makes the selection process in lexical decision tasks quicker and easier for emotional words compared to other types of AWs. The semantic characteristics of these words and the level of emotionality facilitate how speakers with only one language respond to AWs in a time-controlled environment. Even for non-emotional words, the literature hints that these words may still have limited emotional intensity which could impact the processing of these words in the lexicon (Marmolejo-Ramos & Dunn, 2013). Taken together, by being evident in emotional words and limited in non-emotional words, this model addresses the effect of emotionality on the representation of AWs in the lexicon. Overall, the role of emotionality in AWs is the only difference between classical and recent understandings of the representation of these words. Vigliocco and Kousta argued that AWs could not be linked to one static analogue image alone and that the representation of these words in the lexicon goes beyond static images (Kousta *et al.*, 2011; Vigliocco *et al.*, 2011). Research in the late 1990s and the beginning of the 2000s was a turning point in studying these words, especially with the emergence of Vigliocco and Kousta's model later on.

For the current study, whilst the DCM provided a foundation for comparing the patterns of AWs with CWs, its failure to consider the variations inherent in AWs, and the way they are presented in the lexicon. The emotionality model would be more effective in determining patterns in the variation of AWs. The DCM was considered valuable, however, in relation to CWs and charting memory patterns among monolinguals and L2 speakers. Thus, as imagery is a feature of CWs, and emotionality is a feature of AWs, it was necessary to combine these two different models within the study to interpret the emerging patterns appropriately and precisely.

1.3 Key Concepts and Terminology in the Study

In line with the complexity of AWs, there are some vague terms related to the description of the cognitive processes of AWs. It is, therefore, essential to provide a clear list of short and precise explanations of the most related concepts and terms relevant to the research. The following terms are presented alphabetically for the reader’s convenience.

Abstract words. Terminologically, imageless words can be used as a synonym for AWs. This term refers to words with no fixed referents and images. At this point, AWs can be divided into: emotional and non-emotional words (Vigliocco *et al.*, 2009). Others, like Khokhlova (2014), have classified AWs into: sensual and logical words. Although Khokhlova used different terms, he still refers to the same concepts. However, this classification by Khokhlova of non-emotional words as logical words raises concerns. Not all non-emotional words are logical or refer to logical concepts. Therefore, this study will, where necessary, classify AWs as emotional or non-emotional words. Based on this, emotional words are classified into six categories: love, joy, anger, sadness, fear, and surprise (Shaver *et al.*, 1987). These types of AWs contain two extremes of emotional response, either positive or negative. Emotional words have different characteristics because they refer to affective statuses, previous experiences and old memories which differ from non-emotional words in terms of their semantic referent (Altarriba *et al.*, 1999; Mazzuca *et al.*, 2017). Conversely, a word, like *knowledge,* can be understood with no reference to emotion. This example represents a group of words referring to concepts, ideas and beliefs that carry no apparent emotion (Novitskaya, 2006). Most of these words are neutral, concentrating on thought, intellect, faith and logic (Altarriba & Bauer, 2004; Khokhlova, 2014). These words are labelled as AWs because they describe things that cannot be physically observed (Altarriba *et al.*, 1999).

Linguistically, both types of AWs are not restricted to nouns. Some AWs can be adjectives (e.g., *beautiful*), verbs (e.g., *care*) or even adverbs (e.g., *lonely*). Some abstract nouns can be either singular (e.g., *pride*, *prides**) or plural (e.g., *concept*, *concepts*). Moreover, emotional and non-emotional words can be preceded by either definite or indefinite articles. All AWs, regardless of their morphological categories, can also have a metaphorical meaning, depending on the context in which they are used. For example, a positive emotional word, like *hope,* is identified differently when it is used metaphorically (e.g., *hope is on the horizon*), suggesting it is something physically observed by the eyes.

Associations describe the links that connect words in the lexicon (Kiss, 1968; Paivio, 1971a; Wilks & Meara, 2007). There are various terms in this study that refer to the same concept (e.g., *links, associations, connectors, mediators, network,* and *word web*). They can include lexical or conceptual links used for processing L2 words based on the conceptual framework of L1 (Kroll & Stewart, 1994). The characteristics of these links in L2 speakers differ from L1 speakers (Milton, 2009). In psychology, this concept is extended to cover both verbal and visual associations for the same word. These associations refer to any contextual or imaginal links to a particular word. They work as pairs or mediators which can help in memory retrieval.

Conceptual representation refers to how information, ideas, concepts and words are organised in the lexicon according to a set of patterns (Bower, 1970; Hampton & Moss, 2003). All of these genres are stored to be retrieved when needed.

Context availability emphasises the availability of information from the speaker's prior knowledge to explain the meaning of a word. This knowledge is broad, and it explains the linguistic meanings of a word. It is varied in terms of the speaker's linguistic background for a word and the familiarity of using it in a different context. It is also influenced by the speaker's ability to remember words when needed (Field, 2004; Schwanenflugel, 1991; Schwanenflugel & Shoben, 1983; Schwanenflugel & Stowe, 1989).

The emotionality of words refers to the emotional experiences or situations evoked from words (Garrido & Prada, 2018; Kousta *et al*., 2011;

Vigliocco *et al.*, 2013). This form of embodied knowledge is used to process words in the lexicon (Pavlenko, 2008a). The term also describes the levels of intensity reflected by an emotional AW from its positive to negative extremes. Emotionality can also be found among words, as some words are perceived with emotionality, but they do not refer to emotions. It is called, according to Pavlenko (2008a), emotion-laden. For example, a word, like *cancer,* is semantically a CW because it refers to a tangible object, but it can also reflect various levels of negative emotional meaning. Terminologically, the emotionality of words also refers to the emotional intensity or emotional content, and these two terms have been used interchangeably in the literature of AWs.

Imageability refers to the number of images shared with a word or concept. Cognitively, words or concepts with a high level of imageability tend to be accessed and remembered more easily compared to less imageable ones (Altarriba *et al.,* 1999; Field, 2004). Further classification of this term covers analogue and symbolic images representing the word's meaning. An analogue image refers to an image with a direct connection between the word and the image. A word, like *apple,* has different images with different colours, facilitating its meaning. Conversely, symbolic images refer to images with no direct connections that represent the meaning of the word. However, some AWs, like *love,* can be symbolised with some images, like a picture of a heart, which may carry the word's meaning by convention or social attitudes among people. Other AWs, however, do not have any logical connection in linking a word to any imagery. For example, the meaning of the word *pride* can be symbolised by an American flag (Billig, 1995). Images can, therefore, be static, like the image of a heart for the word *love*, or dynamic, like a scene of husband and wife talking to each other.

Words consist of sounds and letters which refer to one or more meanings. Each word is separated from other words when it is used in a sentence (Daller *et al*., 2007). A word can have one letter, like the pronoun I, or more than 28 letters, like *antidisestablishmentarianism*. Words consisting of so many letters are not widely used and, where they are, they have specific contexts. Each word can be further classified into eight groups: nouns, verbs, adjectives, adverbs, pronouns, prepositions, conjunctions and interjections. A word may have one morpheme, like *idea*, or more than one morpheme, e.g., *ideas*. Morphemes are further classified as derivational or inflectional affixes which can attach to the word.

1.4 Structure of the Thesis

The thesis is structured sequentially into eight chapters (including the introduction) to discuss the representation of AWs and obtaining results for L2 speakers. Chapter One has outlined the area of the study and its significance in the field. It aimed to address the gap in knowledge around AWs, and how these words are patterned in the lexicon. By doing this, identifying what is missed in the literature, especially among bilingual speakers becomes clear. This includes introducing research questions, hypotheses and aims. The chapter has indicated the conceptual direction, highlighting models involved in studying AWs. Key terminology has also been addressed to clarify key terms related to AWs. Chapter Two explores the embodied knowledge of AWs. Existing research has shown, for instance, that the representation of AWs in the lexicon is affected by the intensity of emotionality. Emotionality is an essential aspect of embodied knowledge, and it is central to exploring what embodied knowledge is, its features, and which specific parts of these features are applied to AWs. In Chapter Three, the structure of memory, types and characteristics are explored, looking at particular types of memory at the retention stage. Discussing this issue is indispensable for observing how AWs have been examined in the literature.

From here, we move on to three chapters exploring the experimental side of the current research. Each chapter provides an overview of the aims, methods utilised, the selection of AWs and participants, and the findings. Chapter Four, for instance, details how the level of emotional intensity, as one feature of embodied knowledge, was examined in the form of a scale among two different types of speakers, monolinguals and L2 speakers. The chapter models the level of intensity at the word level between emotional and non-emotional words, and this model was examined across the groups. In Chapters Five and Six, steps are taken to measure the recallability of emotional and non-emotional words in STM and LTM. Chapter Five specifically addresses the recallability of emotional and non-emotional words in STM. This part of the research aimed to determine whether or not AWs with a high level of emotionality would be better remembered in STM, and if these differences are the same among monolinguals and L2 speakers. Likewise, Chapter Six investigates patterns of memory in LTM among the same speakers, similarly aiming to measure or model the memory process at the retention stage of emotional and non-emotional words at the LTM testing condition. After that, interpreting and discussing the emerging patterns in Chapters Four, Five and Six is the essence of Chapter Seven. The chapter explains the patterns that emerged for the three experiments, and links these patterns to the representation of AWs in the L2 lexicon. The chapter further includes limitations and contributions of the thesis alongside the implications of the study to the field. Chapter Eight is the final chapter, summarising the thesis' key findings and makes recommendations for future work on AWs. At the end of the chapter, there is a section which answers the research questions stated earlier in Chapter One.

Chapter 2

Background to Study the Embodied Knowledge of Abstract Words among L1 and L2

2.1 Introduction

Over the last three decades, psycholinguists have shown an interest in studying the role of non-linguistic factors, like experiential states, in the organisation of words in the lexicon (Kousta *et al.,* 2009, 2011; Lakoff & Johnson, 1980b; Vigliocco *et al.*, 2010, 2013). They have focused on studying the effect of these factors on the processing of words with various semantic features in the lexicon. In that capacity, language is not only processed merely at the linguistic level, but is also affected by the sensory, experiential or emotionally grounded knowledge built within each word. For example, the lexicon can identify the meaning of a tangible word, like *dog,* aurally by hearing barking or visually by observing the physical existence of the animal. These links to the word *dog* in the lexicon make it easy to determine the meaning of the word and quickly retrieve its image(s). However, recognising the meaning of AWs is based on human emotions and may, therefore, not be fully identified due to a lack of direct images (Kousta *et al.,* 2009). Symbolic visual associations linked to AWs emerge when an AW is seen or heard. For example, seeing a person receiving gifts from his parents or friends could represent the meaning of the word *happy* in the lexicon (Kousta *et al.,* 2011). Alternatively, some other AWs are represented differently according to the emotions evoked, such as the word *death*.

Since studying the embodied knowledge of the word is mainly psycholinguistic in nature, Kousta and Vigliocco (Kousta *et al.*, 2011; Vigliocco *et al.*, 2011) investigated the type of embodied knowledge of words alongside the mechanism for creating this type of knowledge. These interrelated issues represent the essence of this chapter, with a particular reference to AWs. The impact of emotionality, as one model to study the organisation of these words, will be clarified in this chapter. Furthermore, a more detailed description of the concept of the embodied knowledge of words will be provided. The description in this chapter contains various models of embodied knowledge and factors affecting the representation of this knowledge. The last section of the chapter will examine emotionality intensity linked to AWs, drawing on previous studies which have measured the emotionality of AWs among monolinguals and L2 speakers.

2.2 The Concept of Embodied Knowledge

The concept of embodied knowledge is derived primarily from psychology. It combines two words: embodied and knowledge. Starting with the word *knowledge*, this relates to ideas, concepts, faith and realisation about our lived experience (Fuchs, 2012). Part of this knowledge is understood with words which are gained gradually from the early stages of language acquisition and stored permanently in the lexicon (Tulving, 1972). This kind of knowledge is considered to be lexical knowledge which describes how individuals can understand the semantic meaning of the word. It is built into the lexicon, and reflected in the speaker's ability to communicate clearly and comprehensively with others. It can also be extended to cover any contextual links used to explain the meaning of a word at the sentence level where a word might be used metaphorically. A word, like *hand*, for example, refers to the physical part of a body, but at the same time, it can be used to figuratively create an abstract meaning (e.g., 'This is my *hand* in the picture' versus 'Can you give a *hand*?').

The term *embodied* is derived from the speaker's experiences and interactions within different real-life contexts (Rosch & Lloyd, 1978). Over time, a person accumulates many skills, habits and ideas with an increasing number of experiences, memories and emotions. The newly acquired information is centralised in the body. From this standpoint, humans acquire embodied knowledge from different situations (Rosch & Lloyd, 1978). The realisation of this knowledge can be observed differently. It can be seen through the skills used to solve a problem, or an idea needed to understand a specific concept. For example, a surgical doctor must know about medicine, but they must also have enough practical experience to be a skillful surgeon. This knowledge is then realised in the surgeon's ability to perform operations competently.

At the word level, the knowledge embodied within words varies from one type of word to another. The lexicon begins by perceiving a word, including its lexical knowledge. Embodied knowledge about a particular word then begins. One way of observing this type of knowledge is by looking at the imageability of words, which reflects the individual's ability to concretise their sensory knowledge. Alternatively, the reaction to emotionally charged words with an angry facial expression or harsh tone is seen as another feature of the embodied knowledge of words. Some words may also imply potential actions. Consider, for example, the word *knife*, which can be perceived as an object used for cutting items or attacking someone. The latter then suggests a negative emotional meaning. Another word, like *murder,* reflects a negative emotional arousal by carrying out the action of killing. The lexicon then has negative embodied knowledge about this word. As stated in Chapter One, this type of word is called emotion-laden. Depending on the individual's experiences, these examples suggest that words pattern differently in how they are embodied in the lexicon, and link it to the existing linguistic knowledge.

2.3 The Background of Embodied Knowledge

The emergence of the embodied knowledge of words was studied as a separate field in psychology (Rosch & Lloyd, 1978). The background to embodied knowledge is broad, therefore, the scope of the discussion here is limited to the background of the embodied knowledge of words only. This section aims to characterise the knowledge centred around the word in creating mental imagery which could represent the meaning of the words among individuals. From this perspective, Lakoff and Johnson (1980b) argued for the existence of non-linguistic knowledge in the lexicon, as one pattern of this knowledge involves the creation of *image schema* (the organisation of words in the mind according to a set of patterns) to carry the meaning (see Section 2.4.2.1 for details). In addition to the semantic meaning, the lexicon perceives various patterns of embodied information with words that vary in type (e.g., *verbal and visual*) and stores them (Scorolli *et al.,* 2011, 2012). This tendency to link the knowledge in the mind with the perceived words is not a new idea. Evidence from studies in psychology and psycholinguistics has proven the relationship between words and body-object interactions, forming one aspect of embodied knowledge related to language (Altarriba & Bauer, 2004; Altarriba *et al.,* 1999; Siakaluk *et al.,* 2008; Toglia & Battig, 1978). As stated in Chapter One, DCM is formulated with the idea of mental imagery which individuals create to represent the meaning of words (Paivio, 1971a, 1973). However, the characteristics of words include more than just a single picture linked to a word. In addition to the role of imageability, the representation of words is affected by the degree of the speaker's familiarity with the words and the degree of concreteness. The effect of these factors varies among individuals. That being said, words tend to be stored differently. Words with high levels of imageability and familiarity tend to be remembered better in any memory task. These two factors, as well as others, represent the essence of the embodied knowledge of words. To date, many studies have aimed to characterise the embodied knowledge of words, including AWs (Altarriba & Bauer, 2004; Altarriba *et al.,*1999; Kousta *et al.,* 2009, 2011; Vigliocco *et al.*, 2010, 2013), especially as AWs do not represent real life items, but are instead abstracted in our minds as internal emotional states (Barsalou & Wiemer-Hastings, 2005). The complexity found in these words attracted the attention of psycholinguists, like Vigliocco and Kousta, who argued that the representation of AWs is affected by embodied knowledge, including experiential affective information. The speakers' experiences with these words can be described as part of the embodied information about these words (Kousta *et al.,* 2011).

2.4 Models of Embodied Knowledge

Models involved in the embodied cognition of words tend to be categorised into: cognitive and neurolinguistic models. As stated in the previous section, the discussion of embodied knowledge emerged originally in psychology with the work of Rosch and Lloyd (1978), supporting the role of non-linguistic factors in the organisation of words in the lexicon. From this standpoint, efforts have been made by Lakoff (1987) and Lakoff and Johnson (1980b) to characterise this knowledge further and to devise a cognitive framework for the conceptual metaphor and image schema of words to provide a better understanding of the embodied knowledge of metaphorical use of words. The second set of models is neurolinguistic. These aim to explain the process of brain stimulation and neuron activation during the establishment of embodied knowledge. The following section sheds light on these two groups, highlighting how they see words as being represented in the lexicon. The section is important to lay down the discussion on the characteristics of embodied knowledge linked to AWs. The discussion is further subdivided into two parts, starting from models which examined larger linguistic units and moving onto smaller ones. Therefore, models of embodied knowledge will be debated in the following section at the sentence level first, followed by the word level.

2.4.1 Embodied knowledge at the sentential level

2.4.1.1 Conceptual metaphor model

Due to the different levels of meaning found in some sentences, the lexicon can recognise the metaphorical meaning by building a framework. This cognitive framework aims to reform the literal meaning of the source domain to the metaphorical concepts through conceptual mapping (Lakoff & Johnson, 1980a). The new meaning, in this case, the metaphorical one, has its semantic origins in using CWs. Czech (2016), for instance, provides an example within the economic context, stating that "the current economic situation by invoking the economy is a ship (The economy of X is sinking)" (p. 343). In this context, the word *ship* is semantically a CW which metaphorically refers to the *economy*. The *economy* has progressively deteriorated, creating a negative image of this word in the lexicon. The newly created image is referred to as the conceptual frame, where the role of the lexicon is to understand the metaphorical meaning or embodied knowledge emerging from this context and frame the meaning in a proper context. To fully comprehend this conceptual mapping, the lexicon starts by understanding the literal semantic meaning of words. The retrieval of the embodied knowledge of these words comes from the memory to which these words have previously been exposed. After that, a combination of both structures is made to understand the metaphorical meaning of these words. More importantly, the conceptual mapping model quantifies the conceptual metaphors of all fields to examine the process of creating a metaphor at a wider level. Overall, metaphors have been studied in specific contexts, such as political and religious discourses (Maalej, 1990; Tracy, 1979).

Though the model devised by Lakoff and Johnson (1980a) has values, there are some limitations, regarding the application of this model at the sentence level. The model does not narrow down the type of embodied knowledge associated with words. As one lexical unit, the word cannot provide a metaphorical meaning unless contextualised. The role of context, then, is fundamental in determining the correct metaphorical meaning of a word. However, some AWs can still reflect embodied emotional knowledge without context. The word *anger*, for example, reflects unpleasant ideas, attitudes or even actions when it is used without context (Barsalou & Wiemer-Hastings, 2005). This example, as well as others, tells us that not all AWs have the same characteristics. Some of these words may have other meaning besides the dictionary definition when used without a context.

To understand the meaning of the embodied knowledge of words, there are three key factors which can be used: language, background knowledge, and the obtained inference (Glenberg & Robertson, 1999). These factors are vital in speakers fully understanding any abstract idea, especially the metaphorical meaning of the word. These factors, in fact, try to explain the mechanism of establishing embodied knowledge, known as the indexical hypothesis

(Glenberg & Robertson, 1999). It considers the role of background knowledge in forming the meaning of abstract notions. From theory to practice, the indexical hypothesis also examines the evoked meanings of words. Consider questions, like "Can you *drive* a car" versus "Do you have an *idea* of evolution theory?". In the first example, knowledge refers to a tangible object, precisely, driving a car as a tangible object, while in the second sentence, having knowledge of something is intangible. Nevertheless, both sentences embody mental concepts which can be open to the possibility of representing intangible meaning through tangible objects.

2.4.2 Embodied knowledge at the word level

2.4.2.1 Image schema model

This cognitive model combines two different terms: image and schema. Both words refer to something else when they are used separately, but when they are joined together, they create a new concept. Semantically, the term image refers to real-life objects. These images are, either analogue or symbolic (Sternberg, 2006). For schemas, the underlying meaning of this word refers to a particular set of patterns and templates which are used to organise information, ideas and items in the mind. When combined, these terms represent the image schema model which refers to the organisation of knowledge to reflect each idea, notion and word via patterns or schemas. The function of images is then to represent these patterns (Johnson, 1987; Oakley, 2007). Depending on the difficulty of the word, the lexicon creates either direct or complicated image schemas to carry out the intended meaning.

Linking the understanding of words to image schema can be traced back to Bartlett's work (1932) on experimental and social psychology, where she examined the role of schema or prior knowledge in the representation of words. The same concept was discussed in detail, highlighting its mechanism and the underlying structure for establishing images by Johnson (1987). The model of image schema emerged in both cognitive linguistics and cognitive psychology for different reasons. In fact, there was a lack of explanation in establishing conceptual metaphor by moving from the literal to metaphorical meaning. The relation between these two levels of meaning and their visualisation to identify the image created in the lexicon necessities explanation. In fact, the example of 'sinking ship' in Section 2.4.1.1 represents a failing economy. In this example, the lexicon is stimulated to retrieve information about previously stored words. The retrieved information is linked to the picture of a failing ship (Cook, 1989). Most debates in cognitive linguistics have considered the process of metaphorical mapping by creating different semantic nodes and networks (Rhee, 2002). In cognitive psychology, there is a need to know how this kind of knowledge is organised according to a set of schemas and images. It is suggested that the relationship between the mind and language is constrained by various patterns involved in forming our sensory perceptual experiences.

Here, the term image refers to tangible things, and encompasses more dimensions of embodied knowledge as templates for formulating our cognition. For example, by observing how the legs of the table provide support for the table, the lexicon identifies the embodied knowledge of the word *support* (Hedblom *et al.,* 2016). This image is used to represent the meaning of words. Therefore, a word will have an image(s) in the lexicon, and these images are stored as they perceive. Another type of image is symbolic which signifies the concept or words being said. They are different from the previous set of words in the sense that these words do not have a real-life representation. A good example here in this context is AWs, and how they are connected to symbolic images. More detail on AWs will be provided later in this chapter.

2.4.2.2 Simulation of words model

The embodiment of words can also be examined from a neurolinguistic perspective. The rationale behind using neurolinguistic tools is based on the fact that the brain tends to link images, sounds, situations and experiences with words reflecting their meanings. When visual or auditory inputs are linked with words, the mind creates embodied knowledge based on our previous experiences. The mechanism for storing words, in particular, is grounded in embodied situations. This view was originally developed by Barsalou (1999) and Wiemer-Hastings (1998). This model in neurolinguistics underpins the concept of verification of the schema of words that have already been stored, and confronted situation or experience. The brain then determines whether the two levels reflect the word's meaning by showing different neural activations for words with more embodied knowledge[5].

2.5 Factors Affecting the Embodied Knowledge of Abstract Words

2.5.1 The availability of context

One factor affecting AWs is the availability of context. This refers to the degree of effort with which the lexicon can contextualise words with a correct sentence, and the ease of finding a context for a particular word (Schwanenflugel, 1992, 1991; Schwanenflugel & Shoben, 1983; Schwanenflugel & Stowe, 1989; Schwanenflugel *et al.,* 1988). The impact of this factor is addressed repeatedly within the current thesis with a particular focus on its impact on the representation of words in the lexicon. As for AWs, these words are known for their complexity in terms of their semantic reference. This complexity is extended when other types of words refer metaphorically to abstract meaning, creating two levels of meaning (e.g., "This is my *hand* in the picture" versus "Can you give a *hand*?"). Without a context, it is, in fact, challenging to determine which meaning is intended. After this step, the lexicon concentrates on the knowledge associated with the meaning. Concerning the availability of context, previous studies conclude that words referring to

[5] This section in discussing this model is short since studying AWs from a neurolinguistic standpoint is not the scope of this thesis.

tangible things are favoured over AWs in several cognitive tasks, like lexical decisions (Schwanenflugel & Stowe*,* 1989), word comprehension (Schwanenflugel & Shoben*,* 1983), and memory (Schwanenflugel*,* 1991). This superiority is attributed to the ease of contextualising CWs compared to AWs. Additionally, more features of AWs were suggested by this model which can be observed by comprehending these words properly through context. This process becomes clearer in establishing more semantic associations with AWs (Barsalou & Wiemer-Hastings*,* 2005; Griffiths *et al.,* 2007).

2.5.2 Age of language acquisition

The age at which language is acquired plays a role in determining the right time to assess the effect of emotionality on AWs. Previous work has demonstrated that AWs, including emotional words, are not acquired in the early stages of language acquisition (Davies, 2016). AWs require a high level of language competency. This indicates that age is an important factor in acquiring these words. From this, it is argued that the environment in which these words are tested is significant because learning a language in a natural setting tends to have more power in terms of emotional intensity compared to when it is learned in classroom settings (Harris *et al.,* 2006). It is, therefore, important to consider this as a factor that could affect language competency in acquiring various types of AWs.

2.5.3 Word frequency

Another factor that affects the establishment of embodied knowledge is the frequency of use in communication. The frequent use of words expands the embodied knowledge in the lexicon and speeds up the retrieval process. The role of the lexicon is considered to be passive when using less frequent words (Laufer, 1997; Quinlan & Dyson, 2008). With less frequent words, the lexicon fails to link to new experiential knowledge, suggesting that the lexicon only rely semantic knowledge. This behaviour makes the retrieval process time-consuming compared to frequently used words. This conclusion belongs to all words, making no exception to AWs.

2.6 Limitations of Embodied Knowledge Model

As fruitful as this model is in explaining the behaviours of AWs, it has some boundaries. The first criticism concerns the confusion over specifying the speakers' semantic and cognitive knowledge of AWs. Some views of embodied knowledge see the representation of AWs based on linguistic knowledge at the sentence level only (e.g., *Schwanenflugel's the availability of context*). Although this view as well as others provide valuable insights into word representation, it is not easy to classify embodied models by examining the emotionality of different types of AWs, like the embodied knowledge of emotional and non-emotional words. At the word level, the features of AWs, and whether these words are processed similarly, were not examined until the emergence of Kousta and Vigliocco's research (Kousta *et al.*, 2011; Vigliocco *et al.*, 2011). Prior to this, AWs were generally treated with no differentiation between positive and negative emotional words. Although this model by Kousta and Vigliocco's was devised more recently, there is a noticeable gap in this model, regarding a shift in research from monolinguals to L2 speakers. This aspect of the embodied knowledge of AWs, therefore, remains unidentified, and, as such, it is a considerable oversight of this model. To clarify, the intensity of emotionality of AWs among various levels of L2 speakers with different linguistic abilities speakers has not yet been explored or modelled. It is not clear whether the emotional intensity of AWs will be different between various levels of L2 speakers. This issue could also be identified as missing from this model and viewed as a limitation.

2.7 Measuring the Level of Emotionality of Abstract Words

Chapter One made reference to Kousta and Vigliocco's model (Kousta *et al.*, 2011; Vigliocco *et al.*, 2011) on the level of emotionality linked to various types of AWs (see Section 1.2.6), and it was suggested that AWs reflect a different distribution of emotionality among L2 speakers compared to monolinguals. This section explores this idea further in relation to how emotionality has been measured in the existing literature. It begins by explaining the rationale behind examining the level of emotionality in various types of AWs, before looking at emotionality among monolinguals and then among L2 learners.

2.7.1 The rationale behind measuring emotionality of abstract words

There was an indication in this chapter to the characteristics of emotionality linked to AWs which varied in their features between previous experiences, diverse memories and affective statuses. However, not all AWs have the same distribution of emotionality. This depends on the type of AWs. As expected, emotional words, as one type of AW, have high levels of emotional intensity compared to non-emotional words. This difference is caused by the semantic difference between these two types of AWs. A word, like *joy*, for example, as a positive emotional word, is linked in the lexicon to memories, experiences and situations. Compared to a word, like *thought*, these associations are decreased or even vanish. With more than one language in the lexicon, measuring the emotionality of various types of AWs becomes essential. In theory, the level of emotionality for a particular AW will be different among both monolinguals and L2 speakers. At this point, there are three possible levels of emotionality for AWs. First, a word, like *joy,* will have the same level of emotionality with the same strength among both monolinguals and L2 speakers. Alternatively, emotionality among L2 speakers may be higher or lower than monolinguals. The following section will discuss previous studies on measuring the emotionality of AWs, starting with monolinguals and then L2 speakers, with a particular reference to task design and results.

2.7.2 Abstract words among monolinguals

Siakaluk *et al*. (2016) conducted a study examining the emotional experience of different AWs using a lexical decision task. The main objective of their study was to see whether the effect of emotional experience changed among monolingual speakers in a time-controlled environment. They had previously hypothesised that AWs with a high emotional experience would be processed faster than less AWs with less emotion (Siakaluk *et al*., 2014). They classified AWs from high to low emotional experience, and monolingual speakers had to respond to these words in the form of a word/non-word lexical decision task, using a range of previously selected AWs (Siakaluk *et al*., 2016). Since one of the objectives of the study was to characterise the *reaction time* (RT) to various types of AWs, the methodology was based mainly on a lexical decision task. However, they reached a conclusion, highlighting the effect of emotional experience in processing AWs differently, ranging from AWs with high to less emotional intensity. Emotionality makes the selection process in lexical decisions quick and easy (Siakaluk *et al*., 2016). At the same time, the semantic meaning of these words facilitates how speakers – either monolinguals or L2 speakers – rate these words. The outcomes of their study showed different intensity levels between emotional and non-emotional words. A lack of variation between positive and negative emotional words was also observed.

Prior to the study of Siakaluk *et al*. (2016), Bradley and Lang (1999) were concerned with rating words in terms of emotionality. They believed that words are not only affected by the level of imageability, concreteness and availability of context, but also by the level of emotionality. Bradley and Lang (1999) classified word norms into: valence, arousal and dominance. Valence refers to the degree of pleasantness found in emotional words. Some words in their corpus ranged from highly pleasant to highly unpleasant. The scale ranged from 1-9, where 1 represents the lowest level, and 9 represents the highest level. A word, like *happy,* has a high level of valence (M=8.21), since it underlies a superior emotional intensity compared to a word, like *abuse,* (M=1.80). Arousal is another dimension of scale rating of words which refers to the degree of calmness. Words range from highly calm to highly excited. A word, like *romantic,* has a high level of arousal (M=7.59), whereas a word, like *quiet,* refers to a low level of arousal (M=2.50). In terms of dominance, this norm is related to the control or power exerted by the word, and whether the word reflects something strong or weak. Bradley and Lang (1999) believed that dominance has less influence on the scale rating of words based on emotionality. Therefore, the scale's outcomes received less attention than other two norms; valence and arousal.

By looking at the types of words included in the corpus, various types of AWs have been studied alongside CWs. Arousal and valence, according to

Bradley and Lang, can also be found not only in AWs but also in CWs. More than 600 words have been rated in their corpus on English speakers. This suggests that rating these words was restricted in terms of language and speakers. These ratings used English words only, varying in type between positive, negative, non-emotional words and CWs. For the type of speakers, data was elicited from English speakers, and there was no indication in the corpus whether they were able to speak any other languages in addition to English. The contribution of Bradley and Lang's study is observed in providing a database of English words based on the degree of valence and arousal. Undoubtedly, this is an important contribution in the field.

2.7.3 Abstract words among L2 speakers

As touched on previously in Chapter One, the number of studies on the distribution of emotionality in AWs is inadequate. One of the earliest studies to explore this was Anooshian and Hertel (1994) who were interested in looking at the level of emotionality associated with words using Likert scale (7-point). On this scale, 1 represents the lowest level, and 7 represents the highest level of emotional intensity. The number of AWs presented was 36, divided into: 18 emotional and 18 neutral words. The study involved 36 bilingual subjects who were bilingual in Spanish and English. Spanish was the native language of half of the participants, and English was the native language of the other half. The design of the task involved both languages. The selection of words displayed in the task was random between the two types of AWs. The presentation of words was different, depending on the participants. In the first half of the words, the first list of words was presented in English and the other in Spanish. For the second half of the words, words in the first list were presented in Spanish, and the other half was presented in English. The study involved other tasks, like an incidental memory task. However, the main focus here is on how emotionality was measured. The study concluded that emotionality differs between emotional words (M=5.79) and neutral words (M=2.09). This conclusion was expected, as emotional words carry more emotional intensity than neutral words. However, this study stated that emotionality was language specific among bilingual speakers regardless of whether bilinguals acquired their L2 late or early. Whether words were presented in English or Spanish had no impact on the language in memory, but the only exception was the level of emotionality in every language. Therefore, emotionality was considered to be language specific.

The topic of emotionality among bilingual speakers was further investigated by Ayçiçeği and Harris (2004). There were several differences between Ayçiçeği and Harris's study (2004) and Anooshian and Hertel's study (1994) in terms of word lists and the language of the study. Ayçiçeği and Harris (2004) used a more detailed word list that did not include emotional and neutral words only. They incorporated more types of AWs (both positive and negative emotional words), reprimands (e.g., *shame on you*) and taboo words (e.g., *shit* and *breast*). As for the language of the study, emotionality was measured among native speakers of Turkish learning English. The study involved other tasks, but the focus here was on the patterns of emotionality in the selected words. Words in this study were classified as positive emotional, negative emotional and neutral (16 for each group), including nine taboo words and seven reprimand words. Like Anooshian and Hertel's study (1994), participants were given three types of AWs shown on the screen visually and auditorily. However, for taboo and reprimand words, the presentation of words was different. They were divided into four lists. In each list, taboo words were presented visually in Turkish. They were presented auditorily in English. Since the study aimed to examine the effect of emotionality on the recallability of words, the study concluded that the impact of emotionality was proven to be effective in both the L1 and L2 participants. Both emotional and taboo words showed an advantage in scale rating and memory tasks in Turkish and English. However, the effect of emotionality on memory was observed, leading participants to remember emotional words better than neutral words. The study concluded that the level of emotionality is language specific, as it varied between languages.

With two different conclusions on the level of emotionality in these studies, it was evident that debate surrounded emotional intensity amongst L1 and L2 speakers. The variation of emotionality in two different types of lexicon attracted the attention of Ferré and others who examined various types of AWs in a series of papers (2010, 2013, 2018). Their study in (2010) measured the degree of pleasantness, as one term used to refer to emotionality in the literature. This task required subjects to encode the scale of words in a time-controlled testing environment. Like Ayçiçeği and Harris (2004) and Anooshian and Hertel (1994), Ferré *et al*. (2010) aimed to measure the degree of the emotional intensity of emotional words among proficient bilingual speakers of Spanish and Catalan. The study recruited four groups of speakers. The first group comprised 26 Spanish speakers who were bilingual speakers of Spanish and Catalan, but their L1 was Spanish. The second group was 32 Catalan-Spanish speakers, and their L1 was Catalan. Participants in the third group were 75 Spanish-English speakers who were considered bilinguals in both languages, but were dominant in one of them. The last group comprised 35 bilingual speakers of Spanish-English who had learned English later in life in a classroom setting, unlike the third group. The 36 words selected for the study, including positive, negative and neutral words. They were divided into two experimental lists. Depending on the type of words, those selected words were chosen based on several confounding variables. Since the scope of the study concentrated on emotional and neutral words, valence and arousal were controlled to limit their effect in the scale. For positive emotional words, the degree of valence was (M=7.08) and (M=5.66) for arousal. The degree of valence for negative emotional words differed, as the degree of valence was (M=2.58) and (M=5.70) for arousal. For neutral or non-emotional words, the degree of valence was (M=4.77), whereas the degree of and arousal was (M=4.37). These findings were low compared to the other types of emotional words.

The results Ferré *et al.*'s encoding task (2010) revealed that the level of emotional intensity varied in the three groups with different patterns. First, there was a significant difference in the scale rating of emotional words across groups. Second, positive emotional words were rated higher than negative words, but no differences were observed in the scale rating of these words. In brief, the study concluded that emotional intensity among proficient bilinguals was the same as among monolinguals and true bilinguals. Even the recallability of words, like emotional words, was not affected by certain factors, like language dominance, the context of acquiring the language or the similarity between the two languages (Spanish-Catalan). This conclusion was also confirmed by Ferré *et al.* (2013), affirming that emotionality has the same characteristics in the two languages among bilinguals. The findings in their work were based on different task designs, including encoding task, but both reached the same conclusion. The encoding task did not measure the degree of emotionality but rather the concreteness of words. The reason behind this change in scale rating words was to distract the participants' attention from emotionality as a feature of the words. Instead, they were asked to rate words based on their level of concreteness. This shift was attributed to determine whether emotional words were remembered better than neutral words. Nevertheless, no scale of emotionality was measured in their study since the main focus of the study was the memory patterns of emotionally charged words.

Characterising the emotional content of words has been examined using different approaches, regarding various degrees of concreteness, type of task and language status. The emotional content of words was measured differently in Ferré *et al.'s* study (2018), using closed questions "yes" or "no". Participants were asked to decide whether a word had emotional content or not. They were told that if they felt that the word has emotional content, they should click on "yes". Alternatively, if they believed the word does not reflect emotionality, they should press "no" on the screen. The appearance of each word was time-constrained, so they had to respond quickly. Otherwise, the word disappeared. Clearly, the task aimed to measure the number of words that have been classified with(out) emotionality alongside the RT for clicking on every word. Ferré *et al.* (2018) believed that RT could reflect more patterns about emotional intensity, mainly if it is conducted on different bilingual groups with different language dominance. At the word level, the study used 144 words divided into: AWs and CWs, and AWs were divided into: positive (e.g., *kiss*), negative (e.g., *death*) and neutral (e.g., *chair*) words. These words were selected on fixed degrees in terms of valence and arousal for emotional words. The degrees of concreteness and imageability were another variable to choose between words. These values were adopted from Spanish databases since words were represented in Spanish.

The participants in the affective decision task were 39 Catalan-Spanish bilinguals who were native speakers of Catalan, and had been exposed to Spanish from early childhood, so they were competent in both languages. Stimuli were presented in Catalan and Spanish. The results demonstrated different patterns in terms of language and words. There was no impact of emotionality about the language, as the outcomes of the task did not reflect any difference in the classification of words, either in Catalan or Spanish. Even though participants were native speakers of Catalan, the establishment of emotionality in the L2 matched their native language. According to Ferré *et al.* (2018), it was expected that emotionality in their native language should be stronger since it was the language used day to day from early childhood. However, the outcome of this task did not confirm their speculation.

For the variation of words, the task reflected two finding related to the RT in classifying the emotional content of words and categorising these words precisely. The task suggested that negative emotional words were responded more slowly than both positive emotional words and neutral words. Also, there was an advantage for positive emotional words over negative emotional words.

The level of emotionality for these types of words was found to be higher than the one found for AWs. For categorising words based on emotionality, the task showed that negative words were classified with lower emotional content than positive and neutral words. Since there were other tasks in the study, Ferré *et al.* (2018) concluded that the bilingual's language did not affect the establishment of emotionality.

2.8 Concluding Remarks

In this chapter, we have addressed the question of how the emotionality of AWs has been examined among both monolinguals and bilinguals. At this stage, several issues have emerged, regarding word type, confounding variables and the approach of examining emotionality. It appears from research that consensus is lacking in relation to the classification of words, and what it means to be abstract. Judging by their name, for instance, *neutral words* can be interpreted differently, as AWs with no reference to emotion can be classified as neutral, like the word *knowledge*. This issue indicates the importance of defining and classifying AWs based on clear criteria. The same applies to variables related to each type of AW. For every type of AW, several variables affect the selection of these words. Less control of these variables may affect the outcomes of emotionality. For emotional words, the selection of words is influenced by the degree of valence and arousal found in each AW. The nature of these variables differs when there is a change in the type of AW. Non-emotional words, on the other hand, have another set of variables, like the degree of imageability and concreteness. Words in this category range between high and low in relation to these two variables. The choice of words should consider these two variables, as selecting words with high and low levels in one group may lead to inconsistent patterns of emotionality.

Another issue observed in the literature is characterising emotionality. There are two ways of measuring the distribution of emotionality within AWs. The first is based on asking participants to rate words using Likert scale, either (1-7) or (1-5). 1 in both versions represents words with a high level of emotionality, whereas the numbers 7 and 5 represent the lowest level (depending on the version of the scale). This way of measuring emotional intensity shows the difference between emotional and non-emotional words, and demonstrates the difference between positive and negative emotional words. More patterns of emotionality are found by using Likert scale to show the distribution of emotional intensity in every word. These patterns are essential, especially if there is a need to compare the level of emotional intensity in different bilingual groups. However, this method has a drawback. Participants experience a task effect since they will rate emotional words with a high emotional intensity. This pattern comes with no surprise since participants rely mainly on the semantic features of the words in rating the scale. But the difference between positive and negative emotional words cannot be predicted. From this standpoint, the importance of running this scale emerged in the literature. On the other hand, emotionality can be measured using a yes/no affective decision (e.g., Ferré *et al.,* 2018). This approach is based on deciding whether or not a particular word has emotional content. Except for RT, no further details regarding the intensity are provided by this task design. The inconsistency in characterising emotionality is, in fact, one of the limitations of this task design, especially if there is an aim to compare the level of intensity in every type of AW on the one hand, and between different types of emotional words on the other. In the next chapter, we will explore memory as a cognitive process, highlighting the stages of storing words in the memory and the factors affecting this process. Attention will also be paid to the types of memory at the retention stage, and how AWs have been measured in the literature.

Chapter 3

Memory and Abstract Words

3.1 Introduction

Memory has different functions, and it is responsible for encoding, storing, holding, manipulating and retrieving items (Tulving, 2000). These functions have two dimensions in the lexicon: cognitive and neurological. The cognitive dimension relates to the mental processes involved in storing items in the memory, following a sequence of stages. The neurological dimension concentrates on the biological activations that happen in the brain when storing items in the memory (Okano & Balaban, 2000). It cannot, however, be assumed that all items stored in the memory follow similar procedures. Storing and retrieving items from memory is a complicated process, underpinning how memory stores items and the items stored vary in nature, including – but not limited to – words, digits, images, events, facts and literature. The storage process is influenced by cognitive factors, such as the level of attention (Baddeley, 2000a, 2007). There are other linguistic factors, like bilingualism, as well as the level of language proficiency which affect how L2 words are stored in the memory (Basnight-Brown & Altarriba, 2007; Kroll & Stewart, 1994). This chapter will begin by identifying the mechanism for storing items in the memory. This will be followed by a discussion of the factors affecting the recallability of words. Part of the discussion was to characterise the stages of storing items – including words – in the memory, starting with encoding, retention and retrieval. Since this thesis aims to examine the recallability of AWs at the retention stage, this stage is further subdivided into storing words temporarily in STM and permanently in LTM. The discussion identifies the concepts of STM and LTM as demonstrated mainly in psychology, highlighting their principles and characteristics, and how they have been examined in the literature. The chapter then moves on to talk about AWs, and how they have been examined in previous studies, highlighting the issues raised in the literature. This latter section is significant since it lays the groundwork for examining AWs in the current research.

3.2 Emergence of Memory Studies in Psychology

The interest in examining memory as a cognitive ability emerged primarily in psychology, with a focus on encoding processes and the duration of stored items in the memory (Atkinson & Shiffrin, 1968; Baddeley *et al.*, 1975; Miller, 1956; Paivio, 1971a, 1973; Schwanenflugel & Shoben, 1983). At the encoding stage, items are encoded, either at the verbal or visual level, or both, depending on the characteristics of the item(s). Items encoded at both levels are more likely to be remembered better (Paivio, 1971a, 1973). However, encoding items at one level leads to instability of memory remembrance. Based on the duration of items, memory is divided into two types STM, where items are stored on a temporary basis, and LTM, where items are stored for long durations or permanently. The first type of memory work that emerged in research was on STM. It was shown that this kind of memory is limited to remembering items between seven and eight items, a characteristic of STM often referred to as Miller's magic number (Miller, 1956). More work has been done to characterise storing items in STM and the relationship between STM and LTM (Atkinson & Shiffrin, 1968; Baddeley, 2000a). Subsequent scholarship identified a different type of temporary memory, termed *working memory* (WM) (Baddeley, 2000a, 2012; Cowan, 2008). Although WM also relates to the temporary storage of items in the memory, it is characteristically different from STM in terms of storing and retrieving words[6].

In language studies, memory was examined to see the recallability of words as a way of analysing how words are encoded in the memory among monolinguals and bilinguals. In Chapter One, the encoding process of words in monolingual speakers' memories was shown to involve visual and verbal codes. Studies by Paivio (1971a, 1973, 1991) and Schwanenflugel *et al.* (1988), for instance, examined the recallability of AWs and CWs at the encoding stage. The results indicated that the encoding process occurs at two levels: verbal and visual. The ability to remember CWs more effectively was proven since these words have been coded at two levels. For example, data has been elicited by testing the recallability of AWs and CWs (Paivio, 1971a, 1973; Paivio & Csapo, 1969). Although what was required from participants in each of these testing conditions differed, it was shown that memory patterns in both remembered CWs better than AWs. More details about the recallability of AWs in STM and LTM will be discussed later in this chapter.

3.3 Memory and Lexicon

Storing items in memory involves different processes, depending on the type of items, and whether they have associated sounds pictures or both. These differences influence the way memory stores them (Baddeley, 2000a, 2012; Baddeley & Hitch, 1974). According to Baddeley and Hitch's model (1974),

[6] The focus of the current research is on STM and LTM only as WM is not included in the current study. Since both STM and WM are referring to temporary storage of items in the memory, it is argued that measuring the recallability of AWs in in STM is enough. More work can be done later to measure the recallability of AWs in WM.

the structure of memory relies on three components: central executive, visuospatial sketchpad and phonological loop. These components work in tandem to store items in memory. The phonological loop is responsible for storing items in connection with their associated sounds. This loop holds the item and remembers its sound when needed. The visuospatial sketchpad stores items with visual representation in memory. These two components are controlled by the central executive to classify the function of each component, based on the characteristics of the stored item. All three components represent the essence of storing items in the STM. More emphasis was put on attention in storing items to identify the phonological and visual features of every item to eliminate any possible overlapping with similar item(s) (Baddeley, 2000a, 2012).

Words, as one type of stored item, have visual and auditory information. They have phonological features (e.g., *segmental*) and visual features (*letters*). These features are found in every word when they are stored in STM. For words that are formed using similar letters with only minor differences, the mechanism of storing these words in the memory identifies what commonalities words have in terms of their sounds and spelling (e.g., *adapt* and *adopt, subject* as a verb (V) and *subject* as a noun (N)). Therefore, attention in STM influences the way words are stored in the memory. These mechanisms cover all linguistic components – phonological, morphological, syntactic and semantic – required to store words, particularly in the early stages of language acquisition in STM first and then LTM.

Having set out the characteristics of storing words in the memory, it is important to identify the relation between memory and lexicon. The lexicon is the place for storing large lexical elements (Daller *et al.*, 2007; Pufahl & Samuel, 2014). These elements are ready whenever they are needed in communication. From this, we can argue that memory represents how words are patterned in the lexicon among monolingual or bilingual speakers. The lexicon relies mostly on memory, beginning with recognising the word and then matching it with others in LTM. In any lexical prediction task, the lexicon retrieves words from LTM, especially when predicting the missing words in a context (Ferré *et al.*, 2010). Therefore, the contextual knowledge stored in LTM helps the lexicon to comprehend meaning by retrieving pre-existing knowledge about a particular word. A more detailed discussion on the relation between LTM and the role of context in retrieving words from LTM will be discussed in Chapter Six.

3.4 Stages of Memory Storage

One indication of the complexity of memory is the process of storing items. Unlike its oversimplified view in the classical studies of memory, storing items in the memory involves a complex cognitive process that follows an interconnected sequence of stages, starting with encoding, followed by retention, and finally retrieval (Lockhart, 2000). Each stage is different in terms of its features, and how they should be measured. Identifying these three stages is important to explain at what stage AWs will be examined in the current study.

3.4.1 Memory at the encoding stage

The encoding stage is considered the first stage of storing items in the memory. The aim here is to link the stored items to pre-existing information related to the item (Wegner, 1987). To identify the mechanism of this stage, the encoding process of any word relates to two types of knowledge: linguistic and non-linguistic. As stated earlier in Section 3.3, the linguistic knowledge of a word is divided into small linguistic units which cover phonological (e.g.*, segmental and suprasegmental*) and semantic (e.g., *meaning and orthography*) features. These features are linked to the word in the lexicon whenever a particular word is heard or seen. At the same time, dividing a word into small units helps to exclude other words with similar pronunciation or orthography. Non-linguistic knowledge is linked to memories, stories and experiences related to the word. To link this stage of memory storage to an AW, such as *love,* as an abstract positive emotional word, it is encoded in the memory through both linguistic and non-linguistic knowledge. The linguistic knowledge represents the phonological features of the word /lʌv/ with four letters in order (l-o-v-e). It can be a noun (e.g., they fall in *love* again) or a verb (e.g., my kids *love* chocolate). Semantically, it refers to the passion for liking a person, place, animal or object. Words, like *affection, romance* or *passion,* work as synonyms for the word *love,* whereas words, like *hate,* work as antonyms. Non-linguistic knowledge of the word *love* includes different emotions linked to a mother hugging a child or married couples dining together in a restaurant or analogue pictures.

One of the prominent models that explains the encoding process in memory is Paivio *et al*.'s DCM (1971a). As stated in Section 1.2.6, this model describes how a word is encoded in the memory on both the visual and verbal levels. DCM explains the process of encoding words, suggesting that words do not stand on their own in the lexicon. They are linked with other words, forming networks and associations (Kiss, 1968). The visual and verbal levels work together, where the visual representation concentrates on processing the visual side of the words, and the verbal representation focuses on the linguistic side of them (Paivio *et al.*, 1973). The first level of word representation is the verbal one. This refers to how words are recognised and remembered based on their properties, such as shape and context. Discussing this level of representation was not new in DCM. Morton (1969) emphasised the role of logogens in word representation. Mostly, this view links words to their verbal contextual associations. In DCM, this level refers to the verbal associations linked to words. All words – including AWs – are represented at this level because it represents the linguistic side of the words, including graphemes, phonemes and context. However, according to DCM, AWs are only assumed to be represented at this level since they do not have any physical existence in the real world.

The second level of word representation, according to DCM, focuses on the visual side of words. Recognising a word involves more than contextual links. Imagery, according to Morton's (1969) view, was absent in the structure of the model. Some AWs, for instance, may be linked to images. As stated, a word, like *love,* can be symbolised with a limited static image, like a picture of a heart, which carries the meaning of the word by convention or social attitudes among people. In this example, we can see a justification for linking the word *love* to a picture of a heart or a family social gathering. Other AWs do not have any logical connection to symbolic imagery. For example, though the American flag symbolises the word *pride*, there is no direct connection between the flag and the meaning of the word *pride*. At this point, the characteristics of these words vary, depending on the type of words. Some of these images may have a clear direct relation to the word which could facilitate the meaning of words. Others do not have any link between the word and the image. Further complexity occurs when AWs are encoded in the L2 lexicon, where the intensity imageability of some AWs differs in L1 and L2 speakers. Some AWs, for example, have the same level of imageability, especially those related to emotions. However, the imageability of non-emotional words differs, as these words concentrate on thought, intellect, faith, ideas and logic. The level of imageability of these words also varies between languages. What could be perceived in the speaker's L1 with a high level of imageability is different when a word is said to L2 speakers.

The role of memory, then, is to interplay between the two levels in storing words according to their semantic features. From this, the recallability of words varies accordingly, as some words can be stored visually, verbally or both. Based on a series of memory tasks, following DCM, CWs were remembered better compared to AWs (Paivio & Csapo, 1969; Schwanenflugel & Shoben, 1983; Schwanenflugel & Stowe, 1989). In memory studies, AWs and CWs have been tested in intentional and incidental memory conditions (Paivio & Csapo, 1969). Participants were sorted into groups, based on the type of memory condition. For the intentional group, participants were told from the beginning that they would remember words later, but the incidental group was not. Both AWs and CWs were used for both groups. In both memory conditions, CWs were remembered better than AWs. This seems to support the significance of the duality of representation used in remembering words, like CWs compared with only one level of representation, like AWs.

3.4.2 Memory at the retention stage

In memory studies, the retention stage is the point where items are stored, either STM or LTM (Lockhart, 2000). Based on duration, if items tend to be retained in the memory for 30 seconds or less, they are stored in STM. The number of stored items in STM is limited. If items have been rehearsed many times, they are transferred spontaneously to LTM, where there is no limit to stored items, and they can remain in the memory indefinitely (Cowan, 2008). The classification of the item's behaviours in the memory is based on how long it is retained in the memory. This classification has been widespread in psychology, and it will be discussed in further detail below (see also Chapters Five and Six).

3.4.3 Memory at the retrieval stage

The last stage of studying memory characterises processing words in terms of the retrieval stage or lexical retrieval (Field, 2004). It refers to the mechanism of remembering existing items in memory. This stage, according to Lockhart (2000), is the final phase of memory processes. After encoding and storing items in STM or LTM, this phase retrieves items from memory. It concentrates mainly on the number of retrieved items, and the time required to retrieve these items. There has been a great deal of attention paid to the retrieval stage of memory (Lockhart, 2000). There are four experimental methods to test the retrieval of items from memory: free, cued, serial and prospective. Each of these methods has different characteristics in the way items are retrieved from the memory (Table 1).

***Table 1**. Methods of retrieving words from the memory.*

Methods	*Explanation*
Free retrieval	Words are required to be remembered without providing them with any cues.
Cued retrieval	Words are required to be remembered by providing visual (e.g., *pictures*) and verbal (e.g., *letters*) cues.
Serial retrieval	Words are required to be remembered in order, either backwards or forwards.
Prospective retrieval	Participants are required to remember words and make action(s).

3.5 Factors Affecting Memory

Different factors affect the three stages of memory storage. Some of these factors, however, may either belong to one or two stages of memory storage.

3.5.1 Type of the memory

Whether the speaker is monolingual or bilingual is crucial in how items are encoded in the memory. Since this factor focuses on the number of languages being used, memory encodes words differently. As touched on in Section 1.2.6, Paivio *et al.*'s DCM (1971a) highlights how words are encoded in the memory among monolinguals at the visual and verbal levels. For bilinguals, the process involves different levels of representation when there are two languages in the memory.

3.5.2 Word frequency

Word frequency also affects memory storage. One way of observing its effect is the superiority of remembering the most frequent words compared to the less frequent words. This effect is extended to cover the relationship between storing items in STM and the transference of these words to LTM. Frequent use of an item in temporary storage spontaneously transfers an item from STM to LTM via rehearsal. This indicates that recallability of words is affected whether words have been more frequently used compared to less frequent ones (Hulme *et al.,* 2003).

3.5.3 The availability of context

The work of Schwanenflugel and Shoben (1983) emphasised the role of context in remembering some words. After conducting a series of memory experiments, they concluded that memory retrieval becomes easier and faster if words are provided with a context to explain their meaning(s). As stated in Section 1.3, these contextual associations help in retrieving the required words in comparison to non-contextualised testing conditions. Furthermore, these associations provide, either explicit or implicit knowledge about a word which can be used to predict the required words (Wlotko & Federmeier, 2007). Section 1.3 also touched on the relationship between memory and the lexicon, and how the recallability of words may also depend on the context. Context represents the pre-existing knowledge about a particular word. It helps comprehend the meaning and predict the missing words (Ashby-Davis, 1985; Ferré *et al.,* 2010; Squire, 1987). Therefore, the importance of context in memory studies becomes clear in memory storage, especially in LTM, and the way it should be measured. According to Bogaards (2001), it is assumed that words are not stored in memory alone. During the retention stage in LTM, familiarity with a particular word increases. Speakers may have come across early presented words in isolation or context. This, according to Bogaards (2001), influenced the recallability of these words by creating more associations alongside how the words had been used in context. For some participants, a word, like *ability,* may have created associations, either verbally (in linking to other words) or visually (in linking the meaning of the words to pictures) when it is stored in the memory. This increases the chance of retaining words better and for a longer time.

3.5.4 Word length

Another factor, which affects the recallability of words, is the length of the words, especially in a time-controlled environment, like STM. To clarify the effect of this factor, word length refers to the number of syllables and letters in a word. Remembering words with four syllables or more is, therefore, more difficult compared to words with fewer syllables (Baddeley *et al.*, 1975). The same concept applies to the number of letters in a word (Avons *et al.,* 1994). By comparing the recallability of long words, like *establishment*, to short words, like *nice,* we can observe that short words will be easier to retain and retrieve (Avons *et al.,* 1994).

3.5.5 Level of imageability

One of the word norms found in the literature of memory is related to the intensity of images. Words vary according to the way speakers try to think about possible images. There was an indication of the role of imageability in the memorability of words in Sections 1.2.6 and 1.3. These findings represent the essence of Paivio *et al.*'s DCM in characterising encoding words to images in the memory (1971a). The effect of imageability is also extended to include the retention and retrieval stages too. Imageable words tend to be retained and retrieved better than less imageable ones, and the debate between AWs and CWs highlighted in Section 1.3 is an example (Paivio, 1971a).

3.5.6 Individual differences

In memory studies, individual differences play a role in the way the memory stores words. This is observed in memory studies on the variations in remembering many items which has been shown to be influenced by individuals' age, gender, level of education and left/right handedness (Baddeley, 2000a; Bechara *et al*., 2000). Individual differences also include psychological factors, such as the level of anxiety among speakers, which may impact attention, cause overall deterioration in the recallability of words (Richardson, 1996). According to Paivio (1991), the period between early childhood acquisition and adulthood is particularly significant for word storage. During this period, word storage constantly develops, allowing for the storage of more nonverbal associations. Further differences between young and adult learners are observed in establishing mental imagery, including static and moving images (*Ibid.*).

3.6 Classification of Memory at the Retention Stage

As mentioned above, in memory studies, the retention stage is the point where items are stored, either STM or LTM (Lockhart, 2000). Based on duration, if items tend to be retained in the memory for 30 seconds or less, they are stored in STM. The number of stored items in STM is limited. If items have been rehearsed many times, then they are transferred spontaneously to LTM, where there is no limit in storing items, and they can remain in the memory indefinitely. Therefore, the classification of the item's behaviour in the memory is based on how long it is retained in the memory. This classification has been widespread in the psychology, and will be discussed in more detail below. The following section provides an account of STM and LTM, highlighting the structure and capacity that characterise these two types of memory. The section also reviews related studies which examined the recallability of AWs in these two types of memory.

3.6.1 Short-term memory

Historically, STM refers to what James (1890) called primary memory, where a person needs to safely remember or retain items somewhere in the memory for a limited time during daily activities (e.g., *passwords*). James foregrounded the role of attention in this type of memory because it is limited and remembered for specific moments. Baddeley and Hitch (1974) coined the concept in the 1960s, alongside others, and investigated the concept further in the 1970s, defining it as temporary storage of items in memory. It has a limited capacity in terms of time and items, with capacity referring to the ability of memory to retrieve the stored items immediately or after a short time (Baddeley *et al*., 1975; Cowan *et al.,* 1998).

Using data obtained from a series of memory tasks measuring the capacity of STM, Baddeley *et al*. (1975) and Miller (1956) characterised this type of memory. Starting with the duration of items in STM, this is limited, and can only last for 30 seconds. The capacity of memory ranged between seven to nine items or up to four chunks, known as Miller's magic number. These limitations became evident in the classical studies on STM, whereas later work focused on the order in which items were remembered in an attempt to identify more patterns about STM (Baddeley *et al*., 2009). In fact, there has been some controversy in determining the effect of remembering items in the initial position (primacy effect) or the final position (recency effect). Work by

Baddeley *et al.* (2009), for instance, suggested that items in the initial position are remembered better. They attributed this pattern to the participants' behaviour to keep repeating items in the initial position to avoid forgetting them. As a result, items in the final position are less likely to be remembered. On the other hand, the work of Baddeley and Hitch (1974) and Cowan (2008) stated that there is still a recency effect in remembering items in the final position. They justified this argument based on the fact that items in the final string are still active in the memory, and they will last longer than the ones in the initial position if repeating items in the initial string did not occur.

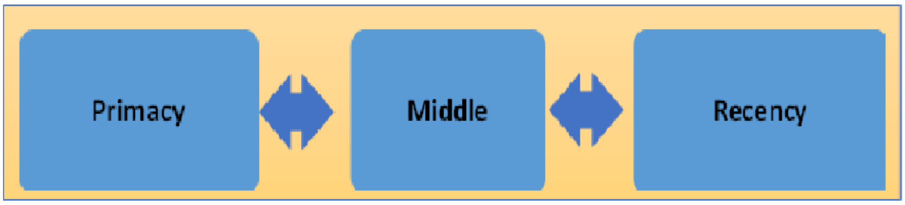

Figure 3. *Distribution of words according to their position in STM.*

At this point, the order of remembrance is an important feature in storing items in STM to determine whether this is a primacy/recency in remembering items in STM (Figure 3). Nevertheless, the markedness primacy or recency effect is affected by the characteristics of the memory task, and the number of items required to be remembered in a timely controlled environment. If the string of items is short (shorter than nine items), there is a chance that the effect of recency will not be observed (Cortis Mack *et al.*, 2017). Alternatively, if the string of items is long (longer than nine items), the recency effect will be clear.

In terms of time, if the time between demonstrating the remembered items, and the time to remember them is long, the effect of recency may not be observed. What can be learned from the order of remembrance is that recallability of items in STM takes different patterns. These patterns are affected by the position of items in the string. This finding is significant in interpreting the results of any memory patterns.

At the theoretical level, one of the well-known models of STM which conceptualised the temporary storage of memory was Baddeley and Hitch's model (1974). The model explains the mechanism of storing words in STM which involves the three components working together to store items temporarily. As stated in Section 3.3, the model structured the mechanism of storing items in STM by dividing STM into components (Figure 4). The central executive coordinates these components to control the capacity to store items in the memory. The model suggests that these components should be available for any stored items, and failure to establish one of these components leads to items being badly retained and an inability to remember them.

Figure 4. *Baddeley and Hitch's model of STM (Source: Baddeley, A.,1986).*

Terminologically, the term STM is confused with WM, because both refer to the temporary storage of items in the memory. WM represents a more complex combination of different elements working together to maintain, manipulate and store items in the memory for a limited period of time. It is defined as the maintenance and controlled manipulation of a limited amount of information before remembrance (Aben *et al.,* 2012; Cowan, 2008). By stating this, further differences between STM and WM are likely to emerge in the way they are tested since WM involves manipulation of the remembered items (Oswald *et al.*, 2015). Therefore, the focus on attention in Baddeley's new version of STM is a key component that allows us to differentiate between STM and WM (Baddeley, 2007). In addition to the temporary storage of items in STM, WM involves more processes to maintain and manipulate the stored items (Cowan, 2008). Taken together, both STM and WM work jointly to store items temporarily which can then lead to them being permanently stored in LTM (Figure 5).

Figure 5. *Baddeley model of WM (Source: Baddeley 2000a).*

This model and Baddeley's view of STM have been criticised for leaving many questions unanswered (Nairne, 2002). As shown previously in Figure 4, the model requires the availability of the three components to store words for a limited period of time. These components are seen as a way of complicating STM, and the way it can be tested while paying no attention to the characteristics of the memory and its limited capacity. The model was a way of visualising the abstract structure of STM only. Yet, this model tried to simplify STM, making it testable under different conditions. Although the model may seem complex, STM can still be tested. Additionally, various factors that can also affect STM were absent in this model. For example, the discussion of frequent items was not discussed. This factor has an impact because frequent items are retrieved better compared to less frequent ones (Alegre & Gordon, 1999). Frequent exposure to items becomes spontaneously part of LTM. The model also did not account for the difference between auditory and visual stimuli, assuming that the stored items should have a phonological loop and a visuospatial sketchpad. To put it differently, the coordination between the components is not clear according to this model. Moreover, some stimuli do not have visual representation, but are still stored in STM (Chan *et al.,* 2008). To explain further, a sound can still be stored in the memory even though there is no visual stimulus representing it. These stimuli are retained in the memory, as the memory attention is high, even though the three memory components are unavailable. Baddeley's model on WM proposed an episodic buffer which combines visual and auditory information and maintains items for a limited period of time (Baddeley, 2000a). Even the emergence of the episodic buffer in this model received less emphasis in discussing the mechanism of storing items in the STM (Nairne, 2002). Baddeley (2000a) subsequently revised his views on STM, and introduced an analysis of the mechanism for storing items in memory. More work by Cowan (2008) tried to describe the mechanism of storing items temporarily from different perspectives. Although these models did not provide a different structure of STM itself, they linked STM to other types of memory and to the role of attention in memory storage. Based on his classic structure of STM, Baddeley (2000a) suggested a revised WM model with a focus on conceptualising attention in storing items temporarily in the memory. STM and WM are considered to be interconnected where there is a link between the stored items in WM and STM since they store items temporarily with a different mechanism (Figure 5). Furthermore, these items are most likely to be stored in LTM (Baddeley *et al.,* 2019; Cowan, 2008).

Overall, these models emphasised the limitations of STM in terms of time and capacity with the relation between STM and LTM.

3.6.1.1 Short-term memory and various types of abstract words

Since the scope of the current study is to concentrate on various types of AWs, this section reviews memory patterns, concerning emotional and non-emotional words. In this section, there will be a detailed discussion on the recallability of emotional and non-emotional words in previous studies among monolingual speakers first, followed by AWs among L2 speakers.

3.6.1.1.1 Emotional and non-emotional words among monolinguals

As indicated in Chapter One, the literature on AWs has not differentiated in memory tasks in measuring the recallability of these words. This is extended to measure the recallability of AWs in STM. It has dealt with all AWs as one group, especially in the classic literature on DCM. There has not been a lot to say about the variation of AWs in DCM literature. This is a key issue in the literature.

Of the limited research in this area, the first task design for measuring the recallability of AWs in STM was by Paivio and Csapo (1969). The design aimed to compare the recallability of these words to CWs and pictures, not only in STM but also in LTM. The reason for this comparison between the three groups was to prove the levels of representation of words which varied between

CWs and AWs, and to determine the retention of these words. Since this section deals with memory patterns of AWs only, details of STM patterns will be included here. Due to the limited capacity of STM, Paivio and Csapo (1969) selected nine AWs for testing conditions, asking participants to remember these words in 30 seconds in an immediate memory task. Items included AWs, like *justice* and *bravery*, CWs, like *piano* and *snake*, and pictures of *piano* and *snake*. Procedurally, the process of measuring the STM of AWs was made in two different testing conditions. Participants had to remember words intentionally and incidentally. In the first condition, participants were told to remember AWs from the beginning. The instruction for remembering AWs was different in the second testing condition, as participants were asked to remember words when the presentation of the words was over. They had no idea before that they had to remember the words which were presented earlier. The outcomes regarding the STM task demonstrated that CWs were better remembered than AWs in the intentional memory condition. The difference between these two types of words was clear, unlike the incidental memory condition. Overall, the memory patterns underlined the significance of imagery in CWs which made the recallability of these words superior. As AWs are represented only at the verbal linguistic level with no or limited visual images, the absence of the visual level, according to DCM, could explain why they are hard to remember in STM. From a broader perspective, Paivio and Csapo's work (1969) did not only provide significant insights into AWs, but it also laid out the principles of measuring STM in terms of the number of words and the time to remember them.

Another task design to measure the recallability of emotional and non-emotional words in particular was based on measuring memory capacity. In this task design, participants (26 monolingual speakers of Spanish) were examined in an encoding task, where they were required to remember as many words as they could (Ferré, 2003). The aim was to examine the recallability of positive, negative and neutral words (20 for each group). The memory task was presented after the encoding task, and participants had to remember words from the previous task. The results showed the superiority of positive emotional words (M=.58) compared to both negative emotional (M=.38) and non-emotional words (M=.08) in a free memory task. At this point, the study used two other memory tasks, but they were cued for perceptual and conceptual cues. To clarify, the perceptual cues were based on providing the initial letters of the studied words. The introduction of letters covered positive, negative and neutral words. These letters served as cues to remind participants about the studied words that needed to be remembered. On the other hand, the conceptual cued task referred to providing participants with a list of words that were associated with the words included in the study, either emotional or neutral. Some of the words were given associates in the list. These words in the list worked as a cue, encouraging participants to remember, and write them down words. For the free memory task, the results showed that positive emotional words were remembered better than negative and neutral words. However, different patterns emerged for both perceptual and conceptual cued tasks. Negative emotional words were remembered better in perceptual cued tasks than positive emotional words. Positive emotional words were better remembered in the conceptual cued task. Neutral words patterned the same in all three memory tasks, as they were remembered the least. Following the same memory task design of Ferré (2003), with a different selection of emotional words, Kensinger (2008) measured the recallability of emotional words with different degrees of arousal. The task design was based on showing five different emotional words, varying in the degree of arousal. They were 15 words taken from five types of emotional words, classified as positive arousing (e.g., *casino*), positive non-arousing (e.g., *lake*), negative arousing (e.g., *slaughter*), negative non-arousing (e.g., *lonely*) and finally neutral (e.g., *figment*). Participants were asked to remember as many words as possible. Each word was on the screen for five seconds. Participants were asked to write down the words they could remember. The results showed that both positive and negative arousing words were better remembered than neutral words among both young and older adults. These results indicated that various types of AWs behaved differently, treating them as one groups was not well-reasoned. The variation between emotional and non-emotional words is an example here.

3.6.1.1.2 Emotional and non-emotional words among L2 speakers

As indicated previously, there has been limited research on memory patterns related to AWs in different memory stages. There have been some select studies, however, which provide important insights into classifications of AWs and the characteristics of memory tasks. Therefore, studies included in this section have been selected based on the classification of AWs and the characteristics of the memory task.

By looking at the literature on these words in an environment where there was less control of time to remember these words, we can say that there was a number of studies which addressed these words among L2 speakers. Anooshian and Hertel (1994) were interested in linking the level of emotionality to memory. They measured the recallability of emotional and non-emotional words through, asking participants to remember as many words as they could from the rating scale task. The number of words included in the scale was 36, varying between emotional and non-emotional words (18 of each). No time frame was applied, however, participants were encouraged to keep remembering words, with no less than 20 words. Those who could not exceed this limit were given more time to remember. The results of this task showed that emotional words were better remembered than non-emotional words. This pattern emerged when the words were presented in the speakers' L1 and not the L2, since the study was conducted on 36 Spanish-English bilinguals.

Ayçiçeği and Harris (2004) followed the same approach in a memory task, asking Turkish-English bilinguals to remember positive, negative, neutral, taboo and reprimand words. After the scale rating of pleasantness, participants were asked as a surprise to remember, as many words as they could in no less than ten minutes. Words presented in the task used Turkish and English in random order. However, it was clarified to the participant that the recallability of words should be in the target language. Words remembered in the participants' L1 were considered wrong. The results indicated that reprimand words achieved the highest level of remembrance, followed by taboo words. Furthermore, positive words were better remembered than negative words.

Looking at whether or not the recallability of emotional and non-emotional words varied between L1 and L2 speakers, as research indicated that remembering emotional and non-emotional words was more effective in the native language (Anooshian & Hertel, 1994; Ferré *et al.,* 2013). These studies aimed to measure the recallability of these words in Spanish-Catalan bilinguals. Designing the memory task in these studies was similar to Ferré's previous study (2003) where participants were asked to remember previously presented words. Participants were asked to remember words in any order and language, and had five minutes to write the remembered words on a sheet of paper. The number of words in the encoding task was 36, divided between Spanish and Catalan words. The results showed the superiority of remembering emotional words compared to non-emotional words in each language. This result in Ferré *et al.* (2013) suggested that emotionality affects the recallability of AWs in any language, either L1 or L2. Since the speakers of the study were Spanish and Catalan bilinguals, the impact of emotionality was clear in both languages.

Having stated the impact of emotional content in L1 and L2, the comparison between emotional and non-emotional words, regarding word concreteness, type of task and language status was made to widen the comparison by Ferré *et al.* later in (2018). The study included affective decision, lexical decision and memory tasks conducted on 39 participants who were Catalan-Spanish bilinguals. They were native Catalan speakers, and they had been exposed to Spanish from early childhood. So, they were competent in both languages. The result of the memory task demonstrated that positive emotional words were remembered better than negative and neutral words.. The findings suggested that the emotional content of AWs led participants to remember these words better in either L1 or L2.

In terms of studies focusing on the effect of primacy/recency in remembering items, there are two groups. The first group of studies examined the effect of primacy/recency as one characteristic of STM. These studies focused on the mechanism of retrieving items, either in the initial or final position better (Baddeley *et al.*, 2009; Cortis Mack *et al.*, 2017). On the other hand, the second group of studies examined the effect of primacy/recency in language studies (Yoo & Kaushanskaya, 2016; Demaree *et al.*, 2004). Since the discussion in this section on emotional and emotional words, only studies on the second group of studies are included here. It is essential to state before discussing these studies that examining primacy/recency effect was not popular in language studies. It is quite common to examine primacy/recency in psychology. For those studies which examined this factor, measuring the effect of this factor has been made by using a list of words with no classification to emotional and non-emotional words (Yoo & Kaushanskaya**,** 2016; Demaree *et al.*, 2004). The study examined the underlying memory structure in remembering words in WM among bilinguals. The aim was to find a relation between the type of memory and the effect of primacy/recency among bilinguals. Methodologically, participants were given a list of words, divided into two groups: one-syllable and two-syllable. For each group, the number of words varied. Each participant was given three groups of words, starting with 10-word, 15-word and 20-word groups, all of which were classified as one-syllable words. The same applied for two syllable words. The distribution of words in each group was different in terms of primacy and recency. For the 10-word list, the first seven words were classified as primacy, whereas the last three words were classified as recency. For the 15-word list, the first 10 words were classified as primacy, whereas the last five words were classified as recency. As for the 20-word list, the first 13 words were classified as primacy, whereas the last seven words were classified as recency. The order of presenting these words was randomised. Since the study was conducted among 20 Korean-English bilinguals, 135 Korean nouns were chosen alongside their English words. No classification of word types was given in the study. By looking at the list being used in the study, each list combined emotional, non-emotional and CWs. Procedurally, it was the job of participants to remember as many words as possible. The recallability of these words was made by recording the participants' voices while they remembered words. The study concluded that the recallability of words is constrained by linguistic knowledge of these words among bilinguals. In reference to primacy/recency effect, the study indicated that the effect of remembering words in the initial position was stronger in the L1, not L2. The effect of recency was also perceived. These two patterns emerged due to the length of the lists. The effect of primacy is likely to be observed with short lists. For longer lists, the effect of recency emerged.

On the other hand, Demaree *et al.* (2004) examined the recallability of positive, negative and neutral words. The approach to examine the recallability of these words was based on presenting the word list to each participant five times. Each list contained 15 words. Words from one to five in the list represented words in the initial position whereas words 10-15 represented a recency effect. The 65 participants were asked to remember as many words as possible in any order. The results indicated that there was a primacy effect for remembering negative emotional words. However, the recallability of positive emotional words reflects a recency effect.

3.6.1.1.3 Observations in the literature in measuring the recallability of abstract words in the short-term memory

After discussing the memory patterns of AWs in the existing work, several issues emerged concerning the classification of AWs, the design of STM tasks and the variables to be measured in STM tasks. The first issue raised in the literature on studying the memory patterns of AWs in STM was the classification of AWs. This issue began in Paivio's work (1971a) on DCM as stated in Chapter One, where he treated AWs as one type. We have seen examples of word lists adapted in STM experiments which combined both emotional and non-emotional words. Consider, for example, the following words: *justice, ability, ego, moral, bravery, amount, theory, freedom* and *grief*. This list represents AWs, but they are not similar semantically in terms of their meaning and cognitively in the way they are represented in the lexicon. Altarriba and Bauer (2004) questioned the generalisation that all AWs are poorly remembered compared to words with tangible references. At this point, AWs remained controversial because they did not have a consistent pattern. They clarified their argument by measuring the level of imageability within AWs. Words, like *bravery* or *ability* tend to have more images, either static or dynamic, compared to AWs, like, *theory* or *amount*. According to Altarriba and

Bauer (2004), imageability is a fundamental factor in the mental process which affects how words are stored in the memory. AWs with a high level of imageability tend to be retained and retrieved better. Therefore, AWs were further classified in their study into: emotional and non-emotional words. Another piece of evidence used by Altarriba and Bauer (2004) to dispute the generalisation of AWs in DCM was the recallability of both types of AWs. Although the experiment has its issues concerning the design as well as the number of words to be remembered, the experiment came to prove the variations between AWs. The participants' recallability was examined, indicating that emotional words were remembered better (M=9.1) than non-emotional words (M=5.70). Moreover, the behaviours of emotional and non-emotional words differed, concerning concreteness and imageability. Among monolinguals, emotional words were rated higher in terms of imageability and concreteness than non-emotional words (Altarriba *et al.,* 1999). Taken together, the recallability of both emotional and non-emotional words is built within the characteristics of each type of AW (Hulme *et al.,* 2003). Based on these findings, AWs varied significantly in their memory patterns, disputing the generalisation of AWs stated in DCM.

Secondly, there was inconsistency in the design of memory tasks in the literature which led to discrepancies in interpretation. By looking at the literature on AWs and memory, we can see that the design of the memory tasks did not specify the type of memory they measured. It is true that there were studies that addressed the recallability of emotional and non-emotional words (Anooshian & Hertel, 1994; Ayçiçeği & Harris, 2004; Demaree *et al.*, 2004; Ferré *et al.*, 2010; Ferré *et al.*, 2018), but they had issues in designing these tasks. The first issue observed was the way these words were tested. Participants in Anooshian & Hertel (1994), Ayçiçeği & Harris (2004) Ferré *et al.* (2010) and Ferré *et al.* (2018) were required to remember words which had been previously presented in the emotionality rating task. The number of words presented in the rating task ranged between 16 and 36 words. The time to remember these words was five minutes, instructing participants to remember as many as possible. By comparing the task design used in these studies to Paivio and Csapo (1969), there were differences in terms of the number of words and the time allotted to remember them. Therefore, it is hard to say that the design of the task used by Anooshian and Hertel (1994), Ayçiçeği and Harris (2004), Ferré *et al.* (2010) and Ferré *et al.* (2018) considered these restrictions in terms of words and time. Therefore, the outcomes of these studies tell us about the recallability of remembering various types of words in the form of free memory tasks, not STM task. Moreover, these studies did not examine variables in measuring the recallability in STM task, like the effect of primacy and recency. Therefore, the design used by these studies could not be implemented to characterise the recallability in STM.

3.6.2 Long-term memory

LTM focuses on storing items for long periods with no limit to its capacity. The role of LTM is to preserve items, and use them when needed. This type of memory was termed secondary memory in the classic work in psychology by James (1890). It is defined as a record of different types of knowledge for a long period of time (Cowan, 2008). It is further subdivided into two main types: declarative and procedural LTM. This classification is based on the characteristics of the stored items and the purpose of storing them in the memory. Declarative LTM represents an explicit memory that deals with the speaker's conscious knowledge of facts, words, concepts and events. It is further subdivided into: semantic and episodic memories (Baddeley *et al.,* 2009. Starting with semantic memory, it is the type of memory which is connected mainly to language. This kind of memory, according to Tulving (1972), is considered to be the mental dictionary which stores words and their associations together in one place. The storing of words is extended to cover the meaning of the words, and how they are combined in a context (Tulving, 1972). Episodic memory relates to remembering events, experiences and dates which are retrieved to create a meaningful context for words. The second type of LTM - procedural LTM - refers to the subconscious knowledge used to perform tasks or skills (Baddeley *et al.,* 2009).

Characteristically, the structure of LTM has been linked to STM, suggesting that the memory patterns of STM could be a preliminary step in transferring items from STM to LTM through rehearsal (Atkinson & Shiffrin, 1968). Despite the age of this model, it remains highly influential, with other models adopting Atkinson and Shiffrin's view on LTM for more than fifty years. To store items in LTM, a speaker encodes what they hear or see concerning the words' meaning(s). The process does not occur temporarily, like STM, and is affected by the rehearsal of items throughout life. This process is realised by the sensory ability which starts by storing items acoustically in STM and then semantically in LTM. STM, as a place for storing items temporarily, helps in storing the same items later in LTM (Cowan, 2008). On the other hand, it has also been suggested that STM and LTM are separate, without any interconnection (see Baddeley & Atkinson, 1974 and, more recently, Norris, 2017). The debate on the relation between STM and LTM is endless in the literature, and each side of the argument has provided evidence from cognitive, linguistic and neurolinguistic standpoints.

Part of explaining the concept of LTM is to state the capacity and the number of items stored in LTM. The process of storing items in LTM begins with encoding an item to its meaning in the memory. Then, when a word is needed, for example, in an ongoing conversation, the retrieval process starts by recognising the required word in the context. These two interconnected processes happen when storing items in STM which leads us to assume that there is a connection between storing items first in STM and then transferring them to LTM. To move items from STM to LTM, they are rehearsed many times (frequently used) so that they are part of the speaker's LTM. Also, it is important to state here that the rehearsal takes many shapes, and it does not mean repeating the words many times only. It can include the availability of context during the retrieval process.

3.6.2.1 Measuring the recallability of abstract words in long-term memory

Historically, measuring the recallability of words in LTM emerged first in the literature on DCM (Paivio & Csapo, 1969), involving comparison of the recallability of AWs and CWs. AWs were considered to be any words that do not have a physical existence in real life, with no further classification of the types of AWs. In addition to words, participants were given pictures of a *piano* and a *snake*, and they were asked to remember them. Participants were asked to remember these words intentionally and incidentally. The work aimed to compare the different memory patterns of these three groups of stimuli in different memory conditions, especially at the retention stage. The task design was based on presenting 72 items to the participants, representing each of the three groups of stimuli. It was the job of the participants to remember them. However, the recallability of these items varied, depending on the time given to remember them. For the LTM task, each item was presented to the participants for less than a second (.63). Participants were given five seconds to write down what they could remember. Part of the study's instructions was to divide participants into two halves. Half of the participants were asked to leave the room and return the next week without being told to remember the words. When they returned, the participants in the room were asked to remember the items unexpectedly. As for the other half, they were told in the first week that they would have to remember the items when they returned the second time. Clearly, the distinction between LTM and STM in their study was considered the time taken to retrieve items, either a week or 30 seconds. According to Paivio and Csapo (1969), the time to remember words was, either delayed (LTM) or immediate (STM). The outcomes of this experiment revealed that both CWs and pictures were better remembered than AWs in both the LTM and STM tasks. The pattern of AWs varied when participants were asked to remember them intentionally and incidentally. AWs were better remembered intentionally when participants were instructed to remember them explicitly in the LTM and STM tasks.

Another experimental design to examine the long-term retention of word learning was the work of Bogaards (2001). The design was based on exposing subjects to two tests held at different times. The first test was held immediately after learning the words (STM), whereas the second was held after a longer period of time (LTM). In the latter, Bogaards held the test after two weeks to see whether the words had been retained in the LTM. The tests included a multiple-choice test, asking participants to choose the correct translation of a Dutch word and an open test where participants were asked to give Dutch translation of a French word. Unlike the previous experimental task design by Paivio and Csapo (1969), this measured word retention of LTM with consideration to context. The first task design by Paivio and Csapo (1969) required remembering the form of the word, including remembering letters and sounds, and the second by Bogaards required more knowledge about the word, including the form and the meaning. In fact, participants had to remember the word, using the context (either translation or definition) to help them remembering the required word. Both designs of LTM tasks agree on the necessity of the learning phase before running the experiment as well as the time of asking participants to remember these words (a week).

On the other hand, there were differences between the two designs in terms of the way words were expected to be remembered by the participants. Paivio and Csapo (1969), for example, asked participants to remember words without providing any cues or reviewing the list of words before examining them. Under these circumstances, such design was similar to measuring the recallability of words in STM except for the time to remember words. The design of Bogaards (2001), however, either provided cues during the examination phase or asked participants to provide a translation for each word. The role of context, as a means of measuring the retention of words in the LTM task suggests that words, including AWs, are not stored in the memory alone. This factor should not be marginalised in measuring LTM.

3.7 Concluding Remarks

Memory - as a cognitive process - underlies certain features and restrictions, with variations between one type of memory and another. Therefore, this is one of the benefits of studying the recallability of various types of AWs, taking into consideration the characteristics of memory types in terms of time and capacity. The review of the recallability of various types of AWs in the existing work indicates the need to further examine the memory patterns of AWs. We have seen that while the literature has contributed to knowledge in a variety of important ways, there is a significant gap in knowledge in:

- The recallability of emotional and non-emotional words at the retention stage, specifically in both STM and LTM; and
- Comparison of the recallability of these types of AWs among two different types of the lexicon.

The objective of measuring the recallability of AWs is to examine the memory patterns of various types of AWs and to determine whether AWs with a high level of emotionality are remembered better when compared with words with lower levels of emotionality. Furthermore, there is an interest in determining whether monolinguals and L2 speakers have common/different patterns in different types of AW. To be more specific, the objective for examining the memory patterns of AWs is to consider the principles of STM and LTM in designing both tasks. The following three chapters provide a detailed overview of the experiments employed in the current study and the subsequent findings. This begins with a chapter which examines the emotional intensity of AWs amongst the participants in the current study, before moving on to Chapters Five and Six which measure the recallability of AWs in STM and LTM respectively.

Chapter 4

Measuring the Emotional Intensity of Abstract Words

4.1 Introduction

As discussed in Chapter Two, Kousta and Vigliocco's model on the level of emotionality stated that AWs reflected a different distribution of emotionality at the word level (2010, 2011, 2013). While there has been extensive research into emotionality among monolingual and bilingual speakers, albeit with some lack of consensus here, there is relatively little investigation into emotionality among L2 speakers. This chapter, therefore, aims to fill this lacuna, and examine the distribution of emotionality for L2 speakers compared with monolinguals. The chapter begins with an explanation and rationale for how emotionality was measured in the current research as influenced by previous studies. This precedes the overview of the experiment and its findings.

4.2 Adapting Scale Rating to Measure the Emotionality of Abstract Words

In the literature on measuring emotionality, the insertion of time was measured in affective lexical decision (Ferré *et al.,* 2018). However, the ultimate aim of this task was to examine the distribution of emotionality between monolinguals and L2 speakers, and to demonstrate the strength between various types of AWs. Therefore, the task design of the scale used in this study was based on Likert scale with no reference to RT in rating words. As stated in Chapter Two, the task design of a scale varies between 1-5 and 1-7, and both designs have been used in previous research. However, except for Ferré *et al.* (2013), the majority of research has employed Likert scale with 1-7 points in the scale rating of emotionality (Altarriba & Bauer, 2004; Altarriba *et al.,* 1999; Anooshian & Hertel, 1994; Ayçiçeği & Harris, 2004). At this point, measuring how quickly participants rate words was not a priority, and it would not characterise the strength of emotionality across different types of AWs. Moreover, the task did not investigate the early stages of acquiring AWs to see which type of AW was rated faster. Since this design has been used in several studies on the literature measuring the emotionality of AWs, it has become more standardised, and this design was, thus, considered more reliable, leading to reliable outcomes. The same design was used to measure other word norms, like imageability and concreteness (Altarriba & Bauer, 2004; Altarriba *et al.,* 1999). Therefore, a decision was made to choose 1-7 Likert scale to measure the emotional intensity of AWs in the current study.

The objective of this task was to measure the variations of emotionality in positive emotional, negative emotional and non-emotional words among L2 speakers. In this experiment, 1 represents the lowest level on the scale and 7 represents the highest level of emotionality. The scale rating aimed to compare the variations found in these words to those of monolinguals. The *independent variables* (IVs) in the experimental groups were the degree of emotionality and the level of language proficiency, whereas the *dependent variable* (DV) was the emotionality elicited from all the groups. For the control group, there was only one IV and one DV.

This part of the study was focused on the first research question highlighted in Chapter One which investigates how L2 speakers of English rate various types of AWs. This led to two sub-questions in this experiment:

1. How do L2 speakers with different levels of language proficiency rate emotional and non-emotional words in terms of emotional intensity? Do they share patterns with monolinguals?

2. What is the strength of emotional intensity in emotional and non-emotional words between competent and modest L2 speakers, and between L2 speakers and monolinguals?

There were two hypotheses for this part of the study:

1. Emotionality varies between emotional and non-emotional words among L2 speakers with different linguistic abilities compared to monolinguals; and
2. The strength of emotionality among L2 speakers is higher when compared with monolinguals since two languages are stored in their lexicon.

4.2.1 Experiment 1

4.2.1.1 Design and materials

This experiment assessed the ability of L2 speakers to rate 32 words based on their level of emotionality. Words referred to emotions and feelings were classified as emotional words here (Tables 2 & 3). These words were further subdivided into: eight positive and eight negative emotional words. The remaining 16 words were non-emotional. All words were obtained from Bradley and Lang's Affective Norms for English Words (1999). There was a restriction to control certain variables since we were comparing a specific set of words. These variables were the degree of valence, arousal and the number of letters for emotional words. For non-emotional words, variables were controlled, like the level of concreteness and the number of letters. Bradley and Lang's corpus (1999) has been widely used in some psycholinguistic studies. Therefore, using such norms was encouraging, and can help in selecting the right emotional words, avoiding any possible criticism for selecting words randomly from the dictionary (Ferré *et al.,* 2013; Garrido & Prada, 2018; Hulme *et al.,* 2003). Starting with emotional words, the degree of valence (ranging from pleasant to unpleasant) and the arousal (ranging from calm to excited) were also obtained from Bradley and Lang's work (1999). Choosing a list of words with exact values of variables was challenging. Specifically, the selection of positive emotional words was constrained with fixed values of valence between 7.50-8.60 (average=8.00) and arousal 5.40-7.67 (average= 6.54) (Table 4). The number of letters for both AWs ranged between three and ten. Therefore, efforts were made to choose words that fell between the two discrete references stated above to create the best environment to examine these words. As for negative emotional words, the level of valence ranged between 1.78 and 3.05 (average=2.33). The arousal of these words was constrained between 4.54 and 6.95 (average=6.20) (Table 4).

Table 2. *Affective characteristics of positive emotional words in scale rating.*

Positive Emotional Words	*Variables*				
	Degree of valence (M)	*Degree of arousal (M)*	*Number of letters*	*Frequency*	*CEFR*
Acceptance	7.98	5.40	10	9080	C1
Comedy	8.37	5.85	6	8844	B1
Joke	8.10	6.74	4	9917	B1
Excitement	7.50	7.67	10	8577	B1
Fun	8.37	7.22	3	11234	B2
Glory	7.55	6.02	5	9199	C1
Joy	8.60	7.22	3	9368	B2
Talent	7.56	6.27	6	9254	B1

Table 3. *Affective characteristics of negative emotional words.*

Negative Emotional Words	*Variables*				
	Degree of valence (M)	*Degree of arousal (M)*	*Number of letters*	*Frequency*	*CEFR*
Crime	2.89	5.41	5	10419	B1
Damage	3.05	5.57	6	10659	B1
Depression	1.85	4.54	10	8948	B2
Hate	2.12	6.95	4	10695	C1
Hurt	1.90	5.85	4	10241	B1
Mistake	2.86	5.18	7	10143	B1
Punishment	2.22	5. 93	10	9134	B2
Tragedy	1.78	6.24	7	8198	B2

Table 4. *The average rate of arousal, valence, and frequency of emotional words.*

Variables	*AWs*	
	Positive emotional words	*negative emotional words*
Arousal	8.00	6.20
Valence	6.54	2.33
Frequency	9434	9804

Unlike emotional words, the selection of confounding variables was different for non-emotional words. Controlled variables with these types of words were concreteness and the number of letters. Concreteness ranged from 3.0 to 3.9 (Table 5) as stated in Altarriba's lists (Altarriba & Bauer, 2004; Altarriba *et al*, 1999). The level of concreteness of non-emotional words averaged 3.31 (Table 6). The number of letters varied between three and nine. No cognate or onomatopoeic words were included between English and Arabic. An effort was made to check whether the English AWs had interlingual homophones between the two languages. The morphological category of words was singular nouns. The current selection of words covered all English letters where possible.

Table 5. *Characteristics of non-emotional words in scale rating.*

Non-emotional Words	*Variables*			
	Concreteness	*Number of letters*	*Frequency*	*CEFR*
Ability	3.6	7	10949	B1
Concept	3.2	7	10337	B2
Deal	3.4	7	11336	B2
Heritage	3.6	8	8880	C2
Intellect	3.1	9	7857	C2
Memory	3.8	6	11652	B1
Now	3.5	3	13304	B2
Legend	3.8	6	9507	B2
Option	3.7	6	10819	B1
Origin	3.3	6	9753	B2
Role	3.2	4	10402	B2
Sequence	3.8	8	11541	C2
Suggestion	3.7	10	9.771	B1
Sight	3.5	8	9570	B2
Time	3.9	4	13578	B1
Total	3.7	5	11162	B1

Table 6. *The average rate of concreteness and frequency of non- emotional words.*

Variables	*Non-emotional words*
Concreteness	3.133
Frequency	10651

At this point, it is vital to state that the list of AWs was selected based on a thematic reason, as each type of AW has different semantic characteristics. Based on this, two issues emerged in this experiment as well as the next ones in the following two chapters. This was attributed to the semantic characteristics of the AWs, and control could not be exercised over all other features of the selected words (such as frequency and the degree of difficulty according to the *common European framework of reference* (CEFR)). The first issue is the frequency of AWs and controlling the selection of these words between fixed values. Unlike other factors, frequency was hard to control for AWs. These words have different semantic characteristics, varying between emotional and non-emotional words. For this reason, it was difficult to control the degree of frequency. According to Milton (2009), it is hard to control when the selection of words is based on a criterion, like the semantic characteristics of words. Therefore, choosing AWs between fixed ranges of frequency was not possible in this experiment. As shown in Table 4, positive emotional words ranged between 8844 and 11234, according to the English Lexicon Project (Balota *et al.,* 2007). The average frequency of these words was 9434 (Table 4). For negative emotional words, frequency ranged from 8198 and 11234, as the average frequency between these words was 9804 (Table 4). Non-emotional words, on the other hand, ranged between 8880 and 11652. The average frequency was 10651 (Tables 5 & 6). These words were chosen from Altarriba *et al.*'s scale-rating experiments (1999) as well as the work of Bleasdale (1987), Chiarello *et al*. (1987), Clore *et al*. (1987), Nelson and Schreiber (1992) and Shaver *et al.* (1987).

The second issue was the level of difficulty. This issue was reviewed according to CEFR (Oxford English Dictionary, 2010). By looking at the words selected in this experiment, we can see some variations, regarding the level of difficulty. Words in this frame are divided into three levels: A, B and C, where C represents words with the highest difficulty level, and A represents the lowest level. For each one of these categories, it is further subdivided into: 1 and 2. Starting with the A category, words in this category are classified into: A1 and A2. Words in the A1 category are more common for beginners, whereas elementary speakers know words in A2. When the level of word difficulty increases as in the B category, we can see that these words are also subdivided into: B1 and B2. Words in the B1 are said to be used by pre-intermediate speakers. B2 words, on the other hand, are more commonly known to intermediate speakers. For the C category, classified as the most challenging words, words are also classified into two groups: C1 and C2. Words in the C1 group are known by upper intermediate speakers, whereas advanced speakers know words in the C2 level. These classifications can be seen from a different perspective. Speakers who can use words at the A1 level are classified as basic users, whereas an independent user is described as an L2 speaker who knows/uses words classified as B words. Words in the C category describe a proficient user[7]. In the current experiment, most positive emotional words varied between B1 and B2. Only *acceptance* and *glory* were classified as C1 according to CEFR (Table 2). For negative emotional words, only the word *hate* was classified as C1. The rest of the list ranged between B1 and B2 (Table 3). For non-emotional words, most of the words were, either B1 or B2. Only three words (*heritage*, *intellect* and *sequence*) belonged to the C level. Having said this, there may be implications for the findings of the experiment as a result of these two issues as will be touched on later in Chapter Seven.

4.2.1.2 Participants

Participants were divided into two groups: experimental and control groups. The experimental group included 60 L2 speakers who were learning English as L2, with Arabic as their L1. They had learned English from formal classroom instruction in Saudi Arabia, where English is the L2. The selection of these two groups was possible since Arabic is the mother tongue of the researcher, and finding a sample with 60 participants required coordination with Saudis living in the UK during the data collection. Recruiting participants whose L1 was not Arabic may have raised concerns about having a cognate between AWs selected in the study and AWs in Arabic. Participants had particular characteristics in terms of the level of language proficiency of English, gender,

[7] This classification is based on the common European framework of reference (CEFR) in linking the level of word difficulty to the level of language proficiency of L2 speakers.

age, education and physical abilities. Starting with the level of language proficiency, the study relied on the *international English language test system* (IELTS) as a standardised test to classify the language level of participants in Britain. In its academic version, this test is classified into band scores from 1-9. According to the British Council, band score 9 reflects the highest level of language ability, whereas band score 1 is the lowest level. The band scores of IELTS characterises the level of language proficiency into absolute beginners (1, 2 and 3), beginners (4 and 5), and advanced speakers (6, 7, 8 and 9). Since the current study chose to study participants whose level of language proficiency was advanced and beginners, the recruiting process selected L2 speakers with 4 or 5 band scores for beginners. For advanced L2 speakers, speakers with 6.5 and above were recruited. Although the increase from 6-6.50 may be seen as minor, this was performed to ensure that the level of all participants was above 6. All of the L2 speakers had taken their IELTS test recently since it is part of their *United Kingdom visa* (UKVIA), and the duration of their graduate degrees varied between 1-2 years for beginners and 3-4 for advanced speakers.

At this point, it was important to ascertain whether the time of taking the test mattered in terms of changing levels of language proficiency among L2 speakers. According to the British Council, the result of IELTS is valid for two years. This led us to assume that the speakers' knowledge of English would be stable since they were living in Britain during their academic programmes. Even if there was a change in their language proficiency, the classification band scores in IELTS would still allow us to assign participants to their actual level. To explain, if a participant scored 6.50 in the IELTS, it is less likely that his level of language proficiency would decline since they had been living in Britain during their academic degree. This indicates that their language was improving, but they would still be classified as advanced speakers since the classification of advanced speakers starts from 6 and goes up to 9. This was the case for PhD students at British universities who were involved in reading and writing throughout their doctoral journey. For beginners, their level of language proficiency at the collection of data was assumed to be the same as at the date of their IELTS. If there was an increase, it would be minor since their stay in Britain was shorter than advanced speakers. For beginners, who showed a low level of language proficiency during the initial conversation, even though they had shown the required score, they were excluded from the study. There was a concern in recruiting them as beginners, but they were considered absolute beginners since they had not practiced their language enough.

For the level of education, participants selected in the study were graduate students at different British universities. At this point, it is important to state that advanced L2 speakers were current PhD students during data collection. The requirement of PhD programmes in the UK at the time of the study was for students to have a 6.50 or above to be admitted. Beginners, on the other hand, were students in the pre-sessional course who aimed to start their master's studies. The medium of instruction in these graduate programmes was English. Therefore, the level of education was homogenous for each group based on their current educational status. Participants in each group had similar educational backgrounds. Advanced L2 speakers had acquired both bachelor and master's degrees, whereas beginners had only one college degree.

In terms of gender and age, the selection of participants was limited to males due to cultural differences. It was hard for a male researcher to recruit such a large number of L2 females. It was anticipated that they would not agree to take part in the experiment in a silent room with a male researcher for almost 15 minutes for cultural reasons. On the other hand, asking another researcher to lead the experiment may raise concerns about the way data should be collected. This could raise practical concerns, including the lack of control that the researcher would have over the collection of data. In terms of the age of the participants, they were aged between 20-30. Due to the availability of participants during data collection in the UK, this age limit was chosen. Age in these studies plays a significant role in the mental representation of the lexicon (Lockhart, 2000). The selection of participants was even more meticulous, considering whether participants were left/right-handed. It has been shown, for instance, that the structure of the brain is different for right and left-handed people. Among right-handed individuals, the language areas are located in the left hemisphere whereas emotion processing occurs in the right hemisphere (Bechara *et al*., 2000). Among left-handers, up to 20% have language areas located in the right hemisphere, while the left holds emotional processing (*Ibid.*). Such a difference may affect the results of the study, and make it difficult to generalise the outcomes for L2 speakers.

The control group included 30 male monolingual speakers of English. The reason behind recruiting a monolingual control group in the study was to examine whether AWs are different or similar between monolinguals and L2 speakers. The monolinguals were native speakers of English, and they lived in Britain. They did not consider themselves to speak any other language competently. The same characteristics in terms of gender, age and physical abilities were applied here too to ensure consistency in the criteria of selecting participants. All of the native speakers were male right-handed speakers aged between 20-30. In terms of level of education, they had all completed a bachelor's degree.

Among both the monolingual and L2 speakers, an initial conversation before conducting the experiments was held. This was performed to check the level of language proficiency, gender, age, education and physical abilities. The advertisement of the study stated these criteria, asking L2 participants to bring their IELTS test when they met the researcher (Appendix A). During the meeting, both monolingual and L2 speakers were asked about their educational background and age. When these criteria were met, they were asked whether they were right-handed or left-handed participants. If any of these criteria were not met, participants were politely declined from the experiment. During the experiment, IELTS scores were checked for L2 speakers. If they showed different language levels, they were also politely excluded from the experiment.

The data collection stage began in January 2019, and ran for just over three months. Participants were recruited from language centres and academic language institutions from different British cities including Brighton, Exeter, Leeds, Newcastle, Sheffield, Southampton and York. There were many L2 speakers for data collection in these cities. The procedure of recruiting participants brought more than 100 participants. Only 90 people fulfilled the requirements.

4.2.1.3 Procedure

The task began by asking the participants to imagine a situation in which the words in Tables 3, 4 and 5 were used. Based on this imagined situation, participants were asked to rate the words on a scale from 1 to 7, regarding positive or negative emotionality, where 1 represented words with no emotion and 7 represented highly emotional words. Values between 3 and 4 indicated that the word has an average level of emotionality. Words were presented without articles (*a*, *an* or *the*). Examples were given before starting the task to make sure that the participants understood what they had to do. Qualtrics was used to ask participants to rate these words (Figure 6). The estimated time to complete this task was five minutes. An internet Wi-Fi router was provided during data collection to protect the data from unexpected network failure.

Figure 6*. A sample of the rating task for the word acceptance.*

4.2.1.4 Ethical considerations

A consent form and an information sheet were distributed to all participants, requesting them to read all of the details carefully before signing the consent form. They were also informed at the beginning that they would be paid £10 for their participation. They were given the choice to, either accept or decline their participation in the study. They were told that even if they did not complete the experiment, they would still be paid for their participation. Following the experiment, all participants were thanked for their time and for participating in this project (Appendix A). They were invited to ask any questions, regarding the study with total confidentiality. The study followed the University of York Data Protection Act (1998) in recruiting participants and storing their data. The ethics committee in the Department of Language and

Linguistic Science approved the data collection for the main data collection round. These considerations were taken into account for this experiment and for the other three experiments in the following two chapters.

4.2.1.5 Sample size

In terms of the sample size (n), the number of participants in each group was 30. The selection of this number was based on a statistical reason. According to central limit theorem, recruiting a sample with the ($n \geq 30$) may help guarantee the normality of the data (Field, 2013). This number is ideal from a statistical point of view, however, there are other factors which could affect the normality of data. Further details will be provided in the following section and in the results section about the normality of data.

4.2.1.6 Approaches to recruit samples

An important step in designing the methodology of the study was to minimise subjectivity in choosing the right sample by recruiting participants based on consistent criteria and well-defined approaches. Without them, the recruited population would not have been representative. According to Langdridge and Hagger-Johnson (2009), there are many ways to control the effect of subjectivity in selecting participants. Since there were three samples in this study, the effect of subjectivity was eliminated using:

- A random sample of monolingual speakers of English. This strategy is considered to be significant in experimental sciences in general (Langdridge

& Hagger-Johnson, 2009). By recruiting speakers from different locations, it was more likely that there was no bias in choosing participants. Every speaker was treated independently, making sure that they met all of the requirements;

- A stratified sample of L2 speakers. In this way, participants could not be recruited randomly because two levels of language proficiency were required beginners and advanced L2 speakers. Having this criterion in mind, it was difficult to sample the L2 speakers randomly without systematically identifying each language level; and

- A snowball sample. This was an additional method for recruiting both native and L2 speakers. This involved participants asking their friends and colleagues with similar characteristics to participate (Langdridge & Hagger-Johnson, 2009).

4.2.1.7 Challenges in data collection

Due to the characteristics of the task design, there were some difficulties when recruiting the participants. Many challenges were encountered as a result of these limitations. The first challenge was recruiting right-handed participants. Finding participants with this feature was challenging for all three groups. After coordinating an opportunity to meet them, unfortunately, some participants had not met the study requirements regarding right-handedness. Others had partial visual or auditory weaknesses which were observed during the initial conversation with the participants prior to the experiment. Therefore, a decision was made to politely apologise to the participants, even if the other experimental measures had been met. Another obstacle emerged between the two L2 groups was the level of language proficiency. The study required a level of language proficiency which should be, either beginners or advanced L2 speakers, according to their IELTS score. Unfortunately, some participants presented themselves as competent L2 speakers but upon further checking their IELTS scores, they were instead classified as beginners. The same difficulty occurred when classifying L2 beginners where their knowledge of English was below the beginner's threshold. Without relying on the IELTS score, recruiting L2 speakers raises concerns. Similarly, recruiting participants in the required age bracket was also challenging. Some potential participants below the age of 20 or over 30 contacted the researcher to take part in the study without paying attention to the age limits set out in the call for participants.

Another challenge was related to the time of year that the data was collected. Because this fell at the time when British universities had a holiday period between the Autumn and Spring terms, there was a limited number of monolingual and L2 speakers, specifically in Yorkshire cities who could be recruited. To finish this phase on time, travelling to different British cities including Brighton, Exeter, Leeds, Newcastle, Sheffield, Southampton and York was necessary to recruit the required number of participants. Travelling to conduct the experiments with these participants was costly in terms of time and budget, especially when an unexpected illness or issues with transport occurred.

4.2.1.8 Data analysis and coding

Statistically, various approaches were used to test the differences between independent samples using the same testing environment. The selection of any statistical test was based on several factors: homogeneity at the level of selected participants, word selection and the characteristics of the data. In any experimental study, it is important to explain the reasons for the selection of tests to increase the power of the results of these tasks. The primary concerns of the current analysis were:

- Each recruited sample was assumed to be homogenous in terms of number, age, gender, right-handedness, education and level of language proficiency;
- The selection of each type of AW in terms of valence, arousal, length, concreteness and frequency was assumed to be homogeneous;
- The independence of each participant did not affect other participants; and
- The characteristics of the data.

The data for emotional and non-emotional words among L2 speakers violated normality. There was some skewness in the data across the word types of AWs. Hence, a decision was made to employ non-parametric tests to confirm the differences between these words. The statistical approach included both descriptive and inferential statistics. Descriptive statistics include means (M), median (Mdn) and standard deviation (SD), whereas inferential statistics are the probability value (*p)*. In experimental studies – including psycholinguistics – the *p* level is (0.05) which has been widely used (Field, 2013). Values greater than (0.05) are considered to be insignificant.

With regards to coding the data, data were transformed via *statistical package for the social sciences* (SPSS) with zero errors. For the scale rating of emotionality, coding data was achieved automatically from Qualtrics to the statistical package for the social sciences SPSS with no human interference. By linking Qualtrics to SPSS, there was no requirement to have multiple coders upload the scale to SPSS. For L2 speakers, beginners and advanced L2 speakers were coded as "1" and "2" respectively, and monolingual speakers were coded as "3". For AWs, the coding was based on the types of AWs, and then the number of words. Emotional words were coded as "Em" followed by the number of words, starting from 1 up to 16 (e.g., *Em_1*). For non-emotional words, words were coded as "Non-em" followed by the number of words (Non-em_17). As stated, responses were coded on the scale of the level of emotionality. Moreover, the internal consistency of the scale was also considered. To do that, the Alpha Cronbach (α) was used to test the internal consistency of every item, with the values ranging from 0.0 - 0.9. Any value of (0.7) or higher is considered acceptable (Kline, 2000; Langdridge & Hagger-Johnson, 2009; Peterson *et al*., 2003). The test provided two values: inter-item and item-total values. Conversely, determining the significance between variables was based on reporting the results of *p*-values. What remains significant is between (.05-.001). These ranges have been extensively used in experimental studies in general. What concerns us at this point are the psycholinguistic studies that consider the results to be significant if they range between these two values.

4.2.1.9 Language of the study

The language used as a tool for conducting the experiments was English. There was a reason for selecting this language. First, the availability of corpuses measuring variables in English, like imageability and the availability of context, was clear. There has been tremendous work in English to measure these factors, emerging in 1978 by Toglia and Battig until recent times. The ultimate aim of these norms was to describe words in terms of these factors. Therefore, choosing words based on fixed values in English is more accessible based on the purpose of the current study. There are different types of corpora where the values of variables are described. Values of words in English emerged in the work of Toglia and Battig (1978). From that time until now, each of these corpora has had its approach to elicit values from participants and the type of variables they controlled. Since the level of emotionality is a language-specific feature, and can vary from one language to another, it requires control of some word norms. The diversity of these norms in English was encouraging to select English as the language for the experiment unlike other languages.

4.2.1.10 Results

This section presents the results for the scale, including both descriptive and inferential statistics. Data were computed and analysed, using SPSS. The results for AWs are presented sequentially, starting with positive, negative and ending with non-emotional words. They are shown here numerically in tables.

At first, the characteristics of the data in this task were examined by identifying the variables as, either IVs or DVs. For the L2 speakers, the type of words and the level of language proficiency were IVs. For the monolingual group, only the type of the words was the IV. The DV was the level of emotionality reported by both the L2 and monolingual speakers. Part of studying the data was to look at the reliability of the scale before the analysis. It was measured using Cronbach's α. This step was taken to confirm that this scale had an acceptable level of internal consistency, and that their outcomes could generate reliable findings. α values from (.70) and above are an acceptable level of internal consistency (Kline, 2000). For L2 speakers, all three types of AWs showed satisfactory α outcomes (Table 7). The internal consistency among beginners was (.77) for positive emotional words and (.83) for negative ones. Non-emotional words had the highest reliability of (.93). More acceptable *α* levels emerged among advanced L2 speakers across all three types of AWs. The results showed differences between emotional negative (.91) and non-emotional words (.86). Still, positive emotional words had the lowest, but still acceptable, value of consistency of (.72). We can observe that positive emotional and negative emotional words among native speakers achieved low yet acceptable *α* results, at (.77) and (.73) respectively. However, non-emotional words touched a high level of internal consistency (.95) (Table 7).

Table 7*. α scale of emotional intensity for the L2 groups and monolinguals (n=90).*

Groups	*AWs Types*	*a*
Beginners	Positive Emotional words	.77
	Negative Emotional words	.83
	Non-Emotional words	.93
Advanced	Positive Emotional words	.72
	Negative Emotional words	.86
	Non-Emotional words	.90
Monolinguals	Positive Emotional words	.77
	Negative Emotional words	.73
	Non-Emotional words	.95

All the α values in Table 7 gave us evidence to assume that the internal consistency of the scale was reliable and could generate reliable data. By looking at α values, we can now look at the means of emotionality across the three groups. Starting with the L2 groups, negative emotional words were rated higher among advanced L2 speakers (M=5.12, SD=1.31) than beginners (M=5.02, SD=1.27). The rate of emotionality decreased slightly in positive emotional words, as the descriptive statistics showed in Table 8. Positive emotional words were rated higher among advanced L2 speakers (M=4.82, SD=1.03) compared to beginners (M=4.72, SD=1.05). Unsurprisingly, the lowest level of emotionality was found in non-emotional words, especially among advanced L2 speakers (M=2.39, SD=1.09). In terms of strength, the emotionality of emotional words tended to be higher among advanced L2 speakers. However, the strength of emotionality of non-emotional words tended to be higher among beginners. The distribution of emotionality demonstrated different patterns among monolinguals. Unlike L2 speakers, negative emotional words among monolinguals were rated higher (M=4.93, SD=.92). Non-emotional words in Table 8 were rated the lowest (M=2.04, SD=1.006).

Table 8. *By-subject descriptive statistics of the scale rating of emotionality among monolinguals and L2 groups (n=90).*

Groups	*AWs Types*	*Mean*	*SD*	*Mdn*
Beginners	Positive Emotional words	4.7232	1.05854	4.9375
	Negative Emotional words	5.0286	1.27957	5.2500
	Non-Emotional words	2.9286	1.49107	2.5313
Advanced	Positive Emotional words	4.8286	1.03153	4.8661
	Negative Emotional words	5.1214	1.31374	5.3750
	Non-Emotional words	2.3967	1.09008	2.1875
Monolinguals	Positive Emotional words	4.6375	.93446	4.7500
	Negative Emotional words	4.9333	.92948	4.8750
	Non-Emotional words	2.0444	1.00616	1.6875

Moving to inferential statistics to see whether these differences between words were significantly different, non-parametric tests were used. Based on the fact that the comparison was between three types of words in one group, Friedman's test (χ^2) was considered the most appropriate test[8]. Starting with L2 beginners, the outcomes of χ^2 showed differences between the three types of AWs. The test revealed a significant difference in scaling these words, χ^2 (2, n=30) =41.153, p<.001. The median indicated that negative emotional words were the highest in emotionality (Mdn = 5.25) compared to positive emotional

[8] Friedman's test is a non-parametric test which is alternative if the data violated the assumption of normality. It is used to test the differences between groups (Field, 2013).

words (Mdn = 4.93). Non-emotional words were rated the lowest on this scale (Mdn = 2.53). A further comparison was conducted to identify the difference between these words. The results of the post hoc comparison showed a significant difference between positive and negative emotional words ($p = .029$). More significant differences emerged between positive emotional and non-emotional words ($p < .001$), and between negative emotional and non-emotional words ($p < .001$). For advanced L2 speakers, the same statistical test was used to examine the differences between these words, χ^2 (2, n=30) =39.176, $p < .001$. The median indicated that negative emotional words were the highest in emotionality (Mdn = 5.37) compared to positive emotional words (Mdn = 4.866). Non-emotional words were rated the lowest on this scale (Mdn = 2.18). As the pairwise comparison results confirmed, there was a significant difference between positive emotional words and non-emotional words ($p < .001$). The same significant result emerged between negative emotional words and non-emotional words ($p < .001$). However, no significant difference was found between positive and negative emotional words among this group of L2 speakers. At this point, similar differences to L2 speakers were found among monolinguals. The outcomes of the χ^2 test demonstrated differences between emotional words and non-emotional words. The test revealed a significant difference in scaling these words, χ^2 (2, n=30) =46.57, $p < .001$. The median indicated that negative emotional words were the highest in emotionality (Mdn = 4.87) compared to positive emotional words (Mdn = 4.75). Non-emotional words were rated the lowest on this scale (Mdn = 1.68). The results of χ^2 demonstrated a significant difference between these words ($p<.001$). The outcomes of the pairwise comparison confirmed that there was a significant difference between positive emotional words and non-emotional words ($p<.001$) on the one hand, and between negative emotional words and non-emotional words on the other ($p<.001$). The lack of difference between positive and negative emotional words was even found among the monolingual group ($p=.053$).

Another dimension of examining the comparison between these words across groups was to examine whether the level of emotionality of each particular type of AW differs among L2 and monolingual speakers. Therefore, a decision was made to use Kruskal-Wallis test, as this test was the most appropriate non-parametric test to examine these differences, attributed to the characteristics of the data as well as the kind of comparison required. The results showed only a significant difference between the three groups in rating non-emotional words χ^2 (2, n=90) =6.84, $p<.001$. The results of the pairwise comparison showed no significant difference between positive or negative emotional words across groups ($p=.743$ & $p=.335$ respectively). For non-emotional words, the pairwise comparison results showed a significant difference between monolingual speakers and L2 beginners ($p=.027$). Even between monolinguals and advanced speakers, there was no significant difference between these words (p=.472). Between L2 beginners and advanced speakers, no difference was found in rating these words based on the level of emotionality (p=.692).

4.3 Summary

In reference to the emotional intensity of AWs, theoretical frameworks from the discipline in psychology assist in explaining the complexity of the lexicon when analysing AWs in a linguistic domain. The psychological analysis results in the current task align with the linguistic analysis of the meaning of these words, as some AWs, especially words referring to emotion, have more emotional responses than others. This experiment concluded that the distribution of emotionality among AWs varied at the word level only. The patterns in this experiment asserted this conclusion not simply among monolinguals but also among L2 speakers with different linguistic abilities. More broadly, it can be argued that the representation of AWs has two levels: emotional and verbal. The emotional level is linked mainly to emotional words, but can also be found among non-emotional words with the least effect. The verbal level, on the other hand, belongs to all AWs according to DCM. AWs in the L2 lexicon have been proven to have the same both levels of representation. The following chapter provides a detailed description of the STM task, elaborating on measuring the recallability of various types of AWs and the possible effect of emotionality on recallability among monolingual and L2 speakers.

Chapter 5

Measuring the Recallability of Abstract words in the Short-term Memory

5.1 Introduction

In Chapter Three, the mechanism of storing items in the memory was touched on, highlighting the stages of storing items – including words – in the memory, starting with encoding, retention and retrieval. Since this thesis focuses on examining the recallability of AWs at the retention stage, this discussion is further subdivided into storing words temporarily and permanently (James, 1890). The current chapter addresses the knowledge gap in examining the recallability of emotional and non-emotional words in STM among both monolinguals and L2 speakers. This includes providing the rationale for the design of the STM task which was influenced by issues raised in previous studies, and lays the groundwork for measuring the recallability of different types of AWs in the current study. A detailed overview of the STM experiment is then provided before the results are presented and discussed.

5.2 Rationale for Measuring the Recallability of Emotional and Non-emotional Words in the Short-term Memory

Chapter Three discussed the recallability of various types of AWs in the literature in terms of characteristics, constraints and limitations. The discussion focused on the characteristics of STM, emphasising the constraints in terms of time to remember (30 seconds) as well as the number (7-9 items). It distinguishes STM from other types of memory, like LTM and WM (Baddeley, 2000b). The literature on the characteristics of STM also underlines the distribution of items in the initial and final positions (Baddeley *et al*., 2009). These characteristics were identified, following the first experiment designed to measure STM by Paivio and Csapo (1969). Given the importance attributed to their work and since it was based on the characteristics of STM - unlike other experimental designs in the literature which focused on the amount of remembered words - a decision was made to adapt Paivio and Csapo's task design in the current study to measure the recallability of emotional and non-emotional words. Memory - as a cognitive process - underlies certain features and restrictions, varying between one type of memory and another. Therefore, this is one of the benefits of studying the recallability of various types of AWs, taking into consideration the characteristics of STM in terms of time and items.

Adapting Paivio and Csapo's (1969) design to measure the recallability of emotional and non-emotional words had various advantages and disadvantages. Starting with the advantages, the task was based on STM principles, as they have been stated mainly in psychological resources (Baddeley, 2009; Baddeley, 2017; Demaree *et al.*, 2004; Paivio, 1971b; Paivio & Csapo,1969). In these sources, there was an emphasis on the duration of STM and the number of items demonstrated during the task. These influenced the current experimental design's focus on examining the impact of primacy and recency on the recallability of these words. As stated earlier, dividing the word string into three groups (initial, middle and final) facilitated measuring which AWs were remembered better in terms of their position. This widens our understanding of the patterns of these words. The elicited patterns tell us how speakers remember words, like AWs, and whether they are able to remember words, either in the initial or final string. These cognitive processes in remembering various AWs open new perspectives in linking STM to language science as a psychological topic on storing words among L2 learners with difference linguistic abilities in particular. Memory represents a significant issue faced by L2 speakers in the learnability of words in L2. For AWs, the design by Paivio and Csapo (1969) assists in providing a better understanding of the memory system when newly learned words are presented to L2 speakers, and whether AWs in the initial, middle or final are remembered differently. This characterises whether AWs have one consistent pattern of remembrance when the time is controlled or not. Alternatively, this design can also uncover other patterns relating to AWs, especially when L2 speakers and monolinguals are examined in one study.

On the contrary, there were some perceived disadvantages, regarding the current task design in terms of time allotted to remember these words (30 seconds) and number of words to be remembered (eight words). These limitations were seen as insufficient to characterise memory patterns. Subjects involved in this testing environment might be under stress in remembering words in a time-constrained manner. It was, therefore, deemed necessary to adapt this design to show consideration of STM principles. The following section gives a detailed description of the task, including the design of the task, the selection of words, and participant information.

5.3 Designing Short-term Memory Task to Measure the Recallability of Abstract Words

The design of the STM task in this experiment aimed to measure the recallability of emotional and non-emotional words, and whether there were similarities and differences between monolinguals and L2 speakers of English. This experiment used a short list of words to be remembered. The design of the

STM task was adapted primarily from psychological studies in STM, but with different word lists that suited the purposes of the current study. The task of the participants was to remember a selection of words presented to them. As a consequence, it was not possible to adapt automated memory tests attributed to the selection of words. In that capacity, the current study used this task design as a lens to measure various types of AWs in STM. Therefore, only the word form was required to be remembered in this experiment. To be specific, the IVs in this task were the amount or the recallability of AWs among L2 groups and the level of language proficiency. The DV, on the other hand, was the number of remembered words for emotional and non-emotional words. For the monolingual group, there was one IV and one DV. The IV was the recallability of AWs and the DV was the number of remembered words. Similar to the scale rating experiment, the monolingual group represents the control group, whereas the L2 groups represent the experimental groups.

Like the previous experiment, various sub-questions framed the STM experiment. These were:

1. When the level of linguistic abilities varies, how do L2 speakers with different linguistic abilities remember emotional and non-emotional words in a timely-controlled environment, like STM?
2. Compared to CWs, how do emotional and non-emotional words differ among monolingual and L2 speakers in STM?

3. Does the order of remembering emotional and non-emotional words differ among L2 speakers by showing, either primacy or recency effect in STM? How is that different compared to monolinguals?

Three main hypotheses underpinned this element of the study:

1. Emotional words are remembered more than non-emotional words among L2 speakers;
2. CWs are remembered better than emotional and non-emotional words;
3. The recallability of emotional and non-emotional words demonstrates a primacy/recency effect in remembering these words among L2 speakers.

5.3.1 Experiment 2

5.3.1.1 Design and materials

This experiment included 16 words which were divided into two groups: emotional and non-emotional words. Each group had the same number of words, which was eight words. Another eight CWs were added to the experiment as a control group of words. They were used to widen the comparison between words. Based on Miller's study stated earlier in Chapter Three, all three types of words had the same number. This number of words was also convenient for dividing the number of remembered words based on their position if participants remembered words in the initial (primacy) or final (recency) string. The string of words here was eight divided into two in the primacy and two in the recency position alongside four in the middle. Each participant was given three-word blocks, representing emotional, non-emotional and CWs with no set order to the sequence in which they were presented.

As for the selection of AWs in this experiment, the same variables, as in the scale rating of emotionality, have been considered here. As stated in the previous chapter, the selection of variables depends on the type of AW. For emotional words, the degree of valence ranged from 2.22 to 8.56 (average= 5.47). For arousal, words selected were constrained at between 4.14 and 7.63 (average= 5.96). The selection of words varied with fixed values of these two variables for emotional words. In terms of the number of letters, emotional words ranged between five and seven letters, as these values represented the highest and lowest number of letters (Table 9). Similar to words on the emotionality scale, word difficulty and frequency remained hard to control. Words selected here were constrained to between B1 and B2, following CEFR (Table 9). The frequency of these words varied between 8073 and 10659 according to the English Lexicon Project (Balota *et al.,*2007). The average frequency of emotional words was 9206 (Table 10).

Table 9. *Affective characteristics of emotional words.*

Emotional Words	*Variables*				
	Degree of valence (M)	*Degree of arousal (M)*	*Number of letters*	*Frequency*	*CEFR*
Anger	2.34	7.63	5	9060	B2
Comfort	8.37	5.85	6	8778	B2
Delight	8.26	5.44	7	8073	B2
Divorce	2.22	6.33	7	8667	B2
Humour	8.56	5.50	5	9455	B2
Passion	8.03	7.26	7	8756	B2
Waste	2.93	4.14	5	10207	B2
Damage	3.05	5.57	6	10659	B1

Table 10. *The average rate of arousal, valence, and frequency of emotional words in the STM task.*

Variables	*Emotional words*
Arousal	*5.96*
Valence	*5.47*
Frequency	*9206*

As for non-emotional words, the controlled variables with these types of words were concreteness and the number of letters. The degree of concreteness ranged between 3.0 to 3.9 as stated in Altarriba's wordlist (1999) (Tables 11 & 12). They averaged 3.32 in terms of the degree of concreteness. Similar to emotional words, the number of letters in non-emotional words ranged between four to eight. The frequency of these words varied between 8334 and 11675 according to the English Lexicon Project. The average frequency was 10516. According CEFR, words being selected here were constrained between B1 and B2 (Table 11).

Table 11. *Characteristics of non-emotional words in the STM task.*

Non-emotional Words	*Variables*			
	Concreteness	*Number of letters*	*Frequency*	*CEFR*
Attitude	3.0	8	9935	B1
Choice	3.7	6	11090	B1
Entry	3.6	5	10451	B1
Guess	3.0	5	11496	B1
Issue	3.5	5	11675	B1
Method	3.2	6	10.816	B1
Obey	3.4	4	8334	B2
Concept	3.2	7	10337	B2

Table 12. *The average rate of concreteness and frequency of non-emotional words in the STM task.*

Variables	*Non-emotional words*
Concreteness	3.32
Frequency	10516

For the control group of words, CWs were selected based on the level of concreteness and the number of letters. The selection followed Altarriba *et al.*'s study (1999) by specifically selecting words with concreteness from 6.2 to 6.4. An effort was made to choose word values where the impact of concreteness was minor between the words (Tables 13 & 14). Compared to non-emotional words, the level of concreteness was higher in CWs due to the characteristics of these words, as they refer to tangible object. The number of letters in CWs ranged between three to six. The frequency of these words ranged between 8793 and 11694 averaged at 9899. Except for the words *eagle* and *thumb*, rated as B2 and B1 respectively, all CWs were rated as A1 or A2 according to CEFR (Table 13). For all three types of words selected in this task, including CWs, it was hard to find words with a fixed level of frequency since these words were selected based on a thematic reason (Table 13). The same applied to the degree of difficulty.

Table 13. *Characteristics of CWs in the STM task.*

CWs	*Variables*			
	Concreteness	*Number of letters*	*Frequency*	*CEFR*
Castle	6.4	6	9361	A2
Dog	6.8	3	10974	A1
Eagle	6.7	5	8831	B2
School	6.1	6	11695	A1
Pen	6.6	3	8975	A1

CWs	*Variables*			
	Concreteness	*Number of letters*	*Frequency*	*CEFR*
Street	6.4	6	11240	A1
Thumb	6.7	5	8793	B1
Garden	6.2	6	9325	A1

Table 14. *The average rate of concreteness and frequency of CWs in STM task.*

Variables	*CWs*
Concreteness	*6.48*
Frequency	*9899*

Each word (including the control words) was presented on the screen for one second in lower-case text. The distribution of words was designed to divide words into three positions: initial, middle and final. The total number of words in each block was eight. They were divided into two words representing the initial and final positions. The remaining four words represented words in the middle. By doing so, the effect of both primacy and recency could be measured. The randomisation was extended by presenting words with no similar initial letters in sequence. Words were presented in the centre of a white screen, using Microsoft PowerPoint. Presenting more than one item on the screen may have affected the attention of the speakers (Aben, 2012; Oswald *et al.,* 2015). Therefore, words were presented one after another and the language of these words was English.

5.3.1.2 Participants

The same participants took part in this experiment as in *Experiment 1 (*see Section 4.2.1.2).

5.3.1.3 Procedure

Mock examples were presented to inform the participants about what was required before the experimental stage commenced to ensure that they fully understood the task. For both the practice and experimental lists, each word was presented separately without any context. Participants were required to promptly remember as many words as they could. They were allotted ten seconds to remember the words belonging to each group. In the case where participants were unable to remember a particular word, they were permitted to leave the space empty. In terms of the location, they were tested individually in a silent room, and the remembered words were handwritten by the participants. They were advised to use a soundproof headset so that they would not hear any noise interfering with the experiment. By doing that, the testing environment encouraged the participants to pay full attention. Additionally, refreshments were provided between sessions with permission to stop the experiments to have a short break. Furthermore, verbal encouragement was also given to the participants before the task so that they could pay more attention to the words being presented.

5.3.1.4 Scoring and data analysis

Scoring memory task was a complicated and time-consuming issue here because this task was conducted on paper. Outcomes were checked by the researcher and another reviewer to make sure that they were consistent. The use of this approach was chosen for STM because the scoring procedure involved human interference and judgement. Data were classified based on the position of words in the list. For the initial position, both emotional and non-emotional words were classified as “primacy_emotional_words” and “primacy _non_emotional_words”. For words in the final position, AWs were rated as “recency_ emotional_words” and “recency_non_emotional_words”. Words that occurred in the middle of the string were classified as “middle_emotional_words” and “middle_non_emotional_words”.

As for the data analysis, the main objective was to ensure that these differences are statistically significant or not. By dividing the number of AWs in each trial into: initial, middle and final positions, the results were restricted to a certain set of values. Each position had a fixed value (0/1/2) so, participants could remember a word or two. Possibly, they could not remember any words. Also, words in the middle string had fixed values (0/1/2/3/4) across the three groups. This pattern in dividing the data affected the normality of data as expected. The histogram confirmed the assumption of abnormality of the data (Figure 7). Although the assumption of abnormality is self-explanatory at this point, a few abnormal histograms were included in the figure to solidify the argument. Reporting the histogram of normality of all three types of AWs in initial, middle and final positions across the three groups gave a lot to report in one figure. As a result, it was difficult to assume that the data were normally distributed, and therefore use a parametric test. The characteristics of the data led us to employ a non-parametric test to study the difference between types of AWs. Similar to the scale rating experiment, the statistical approach included both descriptive and inferential statistics.

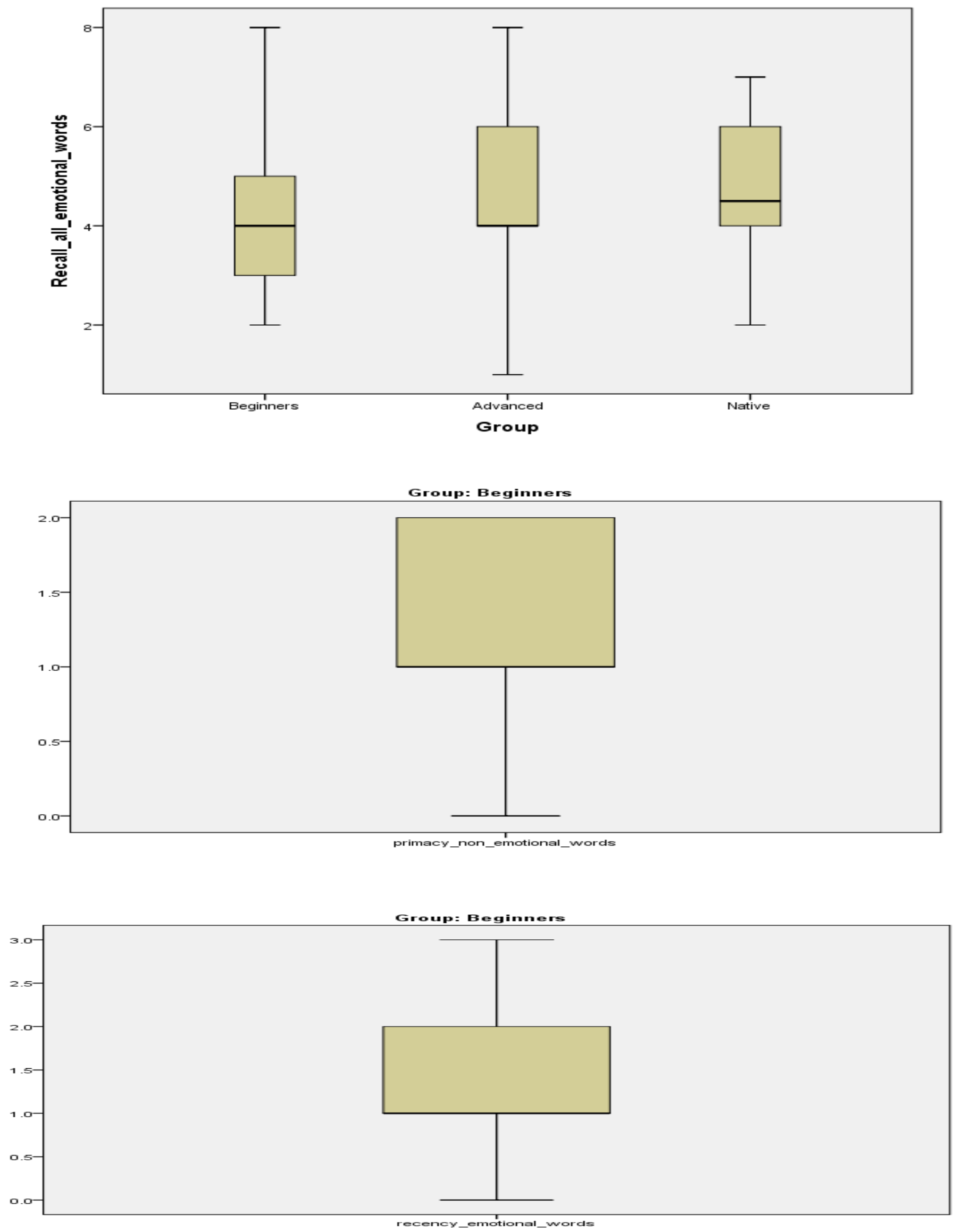

Figure 7. *The characteristics of data in the STM task.*

5.3.1.5 Results

This section demonstrates the results for the STM test, including both descriptive and inferential statistics. The results for AWs are presented sequentially, starting with emotional words and ending with non-emotional words compared to CWs. In this experiment, for L2, there were two IVs: the word type (emotional words, non-emotional words and CWs), and the level of language proficiency (beginners and advanced L2 speakers). The DV was the amount of remembered words, and whether they were remembered better in the initial or final position. The recallability of these words covered the order of remembering words in terms of primacy and recency, and the overall number of remembered words. Unlike the L2 groups, the monolingual group had one IV as there was no level of language proficiency to be examined.

The first step to look at the difference between AWs across the groups was to look at the mean of remembering these words in the initial and final strings. Starting with emotional words, it seems there was a tendency to remember words in the initial string better when compared to the final string among both L2 and monolingual speakers. To specify, the results for advanced L2 speakers (M=1.53, SD=.860) and monolinguals (M=1.50, SD=.509) showed better results in remembering words in the initial string. However, the results for the beginners did not demonstrate any difference between remembering words, either in the initial or final position (M=1.27, SD=.907) and (M=1.30, SD=.702), respectively (Table 15). At this point, advanced L2 speakers and monolinguals behaved similarly, regarding remembering emotional words.

***Table 15**. Results of remembering emotional words depending on the position between L2 and monolingual speakers (n=90).*

Groups	*Order*	*Mean*	*SD*	*Mdn*
Beginners	Primacy	1.27	.907	1.00
	Recency	1.30	.702	1.00
Advanced	Primacy	1.53	.860	2.00
	Recency	1.30	.750	1.00
Monolinguals	Primacy	1.50	.509	1.50
	Recency	.80	.714	1.00

For non-emotional words, the superiority of remembering words in the initial string was observed among monolinguals (M=1.60, SD= .621) compared to beginners (M=1.27, SD=.691). The difference in results between primacy and recency was clear, at least, at the descriptive level. The results for advanced L2 speakers, as stated in Table 16, showed a preference for remembering words in the final string. This was the only unexpected pattern observed in remembering non-emotional words. Compared to the control words, the results for CWs showed that words in the initial string were remembered better among the three groups. Both L2 and monolinguals showed the same pattern (Table 16).

***Table 16**. Results of remembering non-emotional words depending on the position between L2 and monolingual speakers (n=90).*

Groups	*Order*	*Mean*	*SD*	*Mdn*
Beginners	Primacy	1.27	.691	1.00
	Recency	.97	765	1.00
Advanced	Primacy	1.10	.845	1.00
	Recency	1.90	.712	1.00
Monolinguals	Primacy	1.60	.621	2.00
	Recency	.50	.682	0.00

***Table 17**. Results of remembering CWs depending on the position between L2 and monolingual speakers (n=90).*

Groups	*Order*	*Mean*	*SD*	*Mdn*
Beginners	Primacy	1.70	.466	2.00
	Recency	1.20	.714	1.00
Advanced	Primacy	1.70	.535	2.00
	Recency	1.10	.712	1.00
Monolinguals	Primacy	1.92	.305	2.00
	Recency	1.03	.765	1.00

Moving to larger segments, a further comparison was made in this section to examine the recallability of both emotional and non-emotional words by merging words in the initial, middle and final positions. Starting with emotional words among L2 speakers, the descriptive statistics showed that these words were remembered the same across the three groups with only minor differences. Thus, emotional words among advanced L2 speakers were remembered better (M=4.67, SD=4.50) than amongst beginners (M=4.07, SD=4.00). Monolinguals were able to remember emotional words (M=4.63, SD= 4.50) better than beginners. The results for non-emotional words demonstrated low levels of remembrance compared to emotional words. Beginners were able to remember non-emotional words better (M=4.00, SD=1.81) compared to advanced speakers (M=3.70, SD=1.44), but beginners were better than advanced speakers (Table 18). For the control words, all three groups showed a high level of remembrance compared to both emotional and non-emotional words. The results showed that monolinguals had the highest level of remembrance (M=6.00, SD=1.25). Even beginners and advanced speakers, the results showed a high level of remembrance with minor differences (M=5.50, SD=5.50) (M=5.13, SD=5.00), respectively.

***Table 18**. Results of remembering all words in the string of emotional, non-emotional, and CWs (n=90).*

Groups	*Word type*	*Mean*	*SD*	*Mdn*
Beginners	Emotional words	4.07	1.38	4.00
	Non-emotional words	4.00	1.81	4.00
	CWs	5..50	1.45	5.50
Advanced	Emotional words	4.67	1.605	4.50
	Non-emotional words	3.70	1.44	4.00

Groups	*Word type*	*Mean*	*SD*	*Mdn*
	CWs	5.13	1.502	5.00
Monolinguals	Emotional words	4.63	1.42	4.50
	Non-emotional words	3.83	1.23	4.00
	CWs	5.93	1.25	6.00

Differences stated earlier in this section need to be confirmed by inferential statistics. The direction of the analysis has been decided in terms of the use of non-parametric tests to find significant differences between these words. At the word level, a decision was made to examine the difference in remembering emotional and non-emotional words in the initial and final positions among L2 beginners. The test began by measuring the number of remembered AWs in the initial and final positions between emotional and non-emotional words, and whether there was a significant difference. A Wilcoxon signed-rank test (W+) was employed to measure the differences in remembering emotional words. The result of this test indicated that there was no effect of primacy (Mdn, 1.00) or recency (Mdn, 1.00), $z = .057$, $p = .954$. The same pattern of remembrance was also observed among advanced L2 speakers by showing no difference between primacy (Mdn, 1.00) and recency (Mdn, 2.00) for the same type of words, $z = -1.065$, $p = .287$. However, this type of memory pattern of AWs changed among monolingual speakers, showing a primacy effect (Mdn, 1.50) compared to recency (Mdn, 1.00), $z = - 3.33$, p

$<.001$. In terms of remembering emotional words in the middle string, further analysis was conducted to examine the recallability of these words. The results of the Kruskal-Wallis test (χ^2) revealed that there was no difference between the three groups, regarding remembering emotional words χ^2 (2, 90) = 4.411, p=.110. Having stated these patterns about emotional words in three different positions, the total of remembered items was further analysed between the three samples using the Kruskal-Wallis test (χ^2). The results showed no difference in the remembrance of emotional words between the three samples χ^2 (2, 90) = 3.447, p=.178.

As for non-emotional words, W+ was employed. The outcomes of this test gave the same pattern of remembrance, providing no difference between primacy and recency between both levels of L2 speakers. The recallability of beginners did not differ significantly about the primacy effect (Mdn, 1.00), z = 1.33-, p=.181. Likewise, advanced L2 speakers had the same result between primacy (Mdn, 1.00) and recency (Mdn, 1.00), z = - .759, p=.448, indicating no significant difference. Monolingual speakers, however, were able to reflect a consistent pattern between words in terms of the primacy effect (Mdn, 2.00) compared to the recency effect (Mdn, 0), z = - 3.88, p <.001. By moving to a larger unit, the results of the Kruskal-Wallis test (χ^2), showed no significant difference in remembering non-emotional words in the middle string, χ^2 (2, 90) = .732, p=.693 across the three groups. Similar to emotional words, comparing the recallability of all non-emotional words across groups showed no significant difference χ^2 (2, 90) = .447, p=.800.

Having stated these variations of AWs in the experimental groups, the performance of the control words showed a difference between the remembered words in the initial and final positions in all three samples, using W+. Data from beginners showed a primacy effect (Mdn, 2.00) in remembering CWs, z = 2.98-, p =.003. This asserts a primacy effect in remembering these words. The same pattern applied to advanced speakers. The result of W+ showed a primacy effect (Mdn, 2.00) compared to recency (Mdn, 1.00), z = - 2.95, p =.003. For monolingual speakers, the result of W+ demonstrated the primacy effect (Mdn, 2.00) compared to recency (Mdn, 1.00), z = - 4.09, p <.001. For remembering CWs occurring in the middle, the results of Kruskal-Wallis test (χ^2) showed a significant difference between the three samples χ^2 (2, 90) = .6.591, p=.037. The outcomes of the pairwise comparison showed only a significant difference in remembering these words between advanced and monolingual speakers (p=.034). No significant difference was observed between advanced speakers and beginners on the one hand (p=1.00), and beginners and monolinguals on the other (p=.298). The recallability of all CWs across groups showed no significant difference χ^2 (2, 90) = 3.44, p=.100.

5.4 Summary

Examining the recallability of emotional and non-emotional words presented in this chapter has demonstrated different memory patterns. These words behaved among L2 speakers, as they did among monolingual speakers, where there was a reliance on one structure. The results between L1 and L2 speakers in terms of memory provided evidence to prove this assumption. Based on the emerged patterns, the impact of emotionality hypothesised to be reflected in STM was hardly proven under this testing condition. What can be learned about the mechanism of remembering emotional and non-emotional words is that these words have shared memory system in STM when there is a control for time. Such findings widen our understanding of these words, especially if stored among L2 speakers. These patterns will be discussed further in Chapter Seven. Prior to this, the focus of the following chapter is on the recallability of emotional and non-emotional words in LTM, as examined in the current study.

Chapter 6

Measuring the Recallability of Abstract Words in the Long-term Memory

6.1 Introduction

Storing words permanently or in the long-term memory is considered to be another type of memory retention. This is characteristically different from STM in terms of structure and principles as explained in Chapter Three. The current chapter clarifies the rationale for examining the recallability of various AWs in LTM in this study. This step is significant since it lays the groundwork for measuring the recallability of different types of AWs. Next, the chapter discusses two different data collection strategies adopted in the current study for gathering insights on LTM. The second design was a more complex design which featured a learning phase followed by multiple choice and translation tests. Both designs and results are discussed here in detail.

6.2 Measuring the Recallability of Abstract Words in the Long-term Memory

As mentioned above, measuring the recallability of AWs in the LTM involved two different experiments conducted at different stages in the research. The first experiment was based on a cloze test, whereas the second was multiple-choice and translation tests. In this section, further details about each design will be included.

6.2.1 The use of cloze test to measure the recallability of abstract words in the long-term memory

As stated in Chapter 3, the role of context in LTM is significant since it affects the retrieval process of words. Memory is the place to hold these items together. This suggests that measuring LTM requires more cognitive ability to retrieve the correct word from the speaker's LTM. Memory retrieval uses these contextual cues to retrieve the required word to complete the sentence meaningfully. These cues are considered to be part of associations linked to words while storing words in the LTM. The process reflects more than just understanding the sentence and completing what is missing. It characterises the mechanism of retrieving words from LTM.

The recallability of words can be measured by using a cloze test. Characteristically, cloze tests can take many shapes, each has a clear set of objectives fulfilling certain measures. Originally, the main objective of employing a cloze test here was to measure the number of words provided by the participants. Therefore, the focus was on whether speakers were able to assign a particular AW in a given context. At this point, there were two possibilities for designing the cloze test: one *with* and one *without* alternative answers. The former cloze test provides alternative answers which means that the participant can choose the appropriate answer for them. In the latter, they are required to rely on their LTM to remember a word which completes a sentence. The second type, however, enables speakers to remember words and contextualise them. Therefore, the latter was adopted in the initial experiment for the current research.

This type of cloze test has both advantages and disadvantages. Starting with the advantages, adopting this type of design gives the freedom to choose any word that suits the context. No control is imposed on the participant to choose a word previously selected by the researcher. At the same time, it removes any potential for participants to guess based on the options provided, as it is the job of the informants to complete the missing blank by writing a word. This was considered important for the current experiment to measure LTM. Additionally, the cloze test can help in remembering a particular morphological category, since only abstract nouns were examined here. From this, writing sentences which requires only a noun to be completed is essential.

On the other hand, there are some disadvantages to this type of cloze test which are related to the design and the creation of context. The characteristics of this design allow informants to include words which have the same meaning as the expected word, for example, *pleased* instead of *happy*, consequently affecting measurement where synonymous words are used. We have some indication in the literature that synonymous words should be considered (see Chapter Three). The design of the task allows to have various answers. As for the creation of context, contextualising previously selected words is not an easy step. The test items should focus on everyday topics rather than highly sophisticated topics related to science or philosophy. These topics require a deep understanding of the underlying ideas, especially if the sample is not a native speaker of the language (Daller *et al.,* 2007).

6.2.1.1 Experiment 3

The current design for LTM aimed to characterise the memory patterns of emotional and non-emotional words. The aim extended to cover whether there are common grounds between L2 speakers and monolingual speakers of English. Procedurally, the design of this task used a cloze test, where AWs were presented with an accompanying context (Ashby-Davis, 1985; Squire, 1987). The IVs for the monolingual and L2 groups were the recallability of AWs and the level of language proficiency, whereas the DV was the number of emotional and non-emotional words remembered. Similar to the scale rating and STM experiments, the monolingual group represents the control group, whereas the L2 groups represent the experimental groups. In this experiment, essential sub-questions were raised:

1. Similar to the STM task, how do speakers of two languages with different levels of linguistic abilities remember emotional and non-emotional words in LTM? How are they different compared to monolinguals?
2. Compared to CWs, what are the memory patterns of remembering emotional and non-emotional words in LTM?

Similar to the STM task, there were also two hypotheses here:

1. Emotional words are better remembered compared to non-emotional words among L2 speakers in LTM; and
2. CWs are remembered better than emotional and non-emotional words in LTM.

6.2.1.1.1 Design and materials

The experiment included LTM task which required the participants to remember, and write the word that fitted into a sentence. 20 AWs were used during the experiment, and 10 additional CWs were included as a control group of words. The number of words in this task was higher compared to AWs in

STM since the capacity of LTM is unlimited. Starting with AWs, the criteria for selecting emotional words were the same in the scale rating of emotionality and STM tasks in terms of the level of valence, arousal and the number of letters. For positive emotional words, the degree of valance was constrained to between 7.00 and 8.20. As for the level of arousal, it was between 5.20 and 6.74. For negative emotional words, the degree of valence ranged between 1.67 and 6.20. The degree of arousal was constrained at between 4.07 and 6.85 (Tables 19 & 20 show the average of both valence and arousal). For both types of emotional words, the number of letters ranged between four and nine (Tables 19 & 20). By looking at Table 20, we can see that the frequency varied between positive and negative emotional words. For positive emotional words, the words ranged between 8308 and 9917 (average=5725). For negative emotional words, words ranged between 8925 and 9902 (average=9190) according to English Lexicon Project (Balota *et al.*, 2007). The number of letters in negative emotional words ranged between five and eight. In terms of difficulty, both positive and negative emotional words were B2 according to CEFR, except for the words *alert* (C1) and *joke* (B1). Similar to the list of words in Experiments 1 and 2, it was hard to control the level of frequency and the level of difficulty. Words have been selected based on specific semantic characteristics, so controlling these two variables between certain limits was difficult.

Table 19. *Affective characteristics of positive emotional words in the LTM task.*

Positive Emotional Words	*Variables*				
	Degree of valence (M)	*Degree of arousal (M)*	*Number of letters*	*Frequency*	*CEFR*
Promotion	8.20	6.44	9	8787	B2
Joke	8.10	6.74	4	9917	B1
Liberty	7.98	5.60	7	9749	B2
Enjoyment	7.80	5.20	9	8308	B2
Pride	7.00	5.83	5	9015	B2

Table 20*. Affective characteristics of negative emotional words in the LTM task.*

Negative Emotional Words	*Variables*				
	Degree of valence (M)	*Degree of arousal (M)*	*Number of letters*	*Frequency*	*CEFR*
Alert	6.20	6.85	5	9103	C1
Disaster	1.73	6.33	8	9007	B2
Devil	2.21	6.07	5	9016	B2
Fault	3.43	4.07	5	9902	B2
Poverty	1.67	4.87	7	8925	B2

Table 21. *The average rate of arousal, valence, and frequency for positive and negative emotional words in LTM task.*

Variables	*AWs*	
	Positive emotional words	*Negative emotional words*
Arousal	*5.96*	*5.63*
Valence	*7.81*	*3.04*

Variables	*AWs*	
	Positive emotional words	*Negative emotional words*
Frequency	*5725*	*9190*

For the non-emotional words, the controlled variables with these types of words were the degree of concreteness and the number of letters. The degree of concreteness ranged between 3.3-3.8 as stated in Altarriba's list (1999) (Tables 23 & 24). Similar to emotional words, the number of letters in non-emotional words ranged between four and eight. These words ranged in frequency between 9863 and 12174 (average=10825) (Table 23). Words being selected here were constrained to between B1 and B2, following the CEFR of word difficulty (Table 22).

Table 22. *Characteristics of non-emotional words in the LTM task.*

Non-Emotional Words	*Variables*			
	Concreteness	*Number of letters*	*Frequency*	*CEFR*
Character	3.8	9	11062	B2
Culture	3.4	7	10554	B1
Fiction	3.6	7	9863	B1
Knowledge	3.6	9	11338	B1
Mind	3.3	4	11784	B1
Mastery	3.4	7	7198	B2
Opinion	3.7	7	11138	B2
Order	3.4	5	12174	B1

Non-Emotional Words	*Variables*			
	Concreteness	*Number of letters*	*Frequency*	*CEFR*
Position	3.5	8	11884	B1
Response	3.7	8	11264	B2

Table 23. *Average rate of concreteness and frequency of non-emotional words in the LTM task.*

Variables	*Non-emotional words*
Concreteness	3.55
Frequency	10825

For the CWs, these words were chosen based on two factors: the level of concreteness and the number of letters. Words were selected from Altarriba *et al.* (1999) by selecting words from 6.0 - 6.5. An effort was made to choose values for these words where the impact of concreteness was minor between the words (Tables 24 & 25). Compared to non-emotional words, the level of concreteness was higher in CWs due to the semantic characteristics of these words. The number of letters in CWs ranged between four and eight. In terms of frequency, the words ranged between 12174 and 11695 (average=10079). In terms of the level of difficulty, only the word *jewel* was B2, but all the other CWs ranged from A1 to A2. This level is the lowest level of difficulty according to CEFR (Table 24).

***Table 24**. Characteristics of CWs s in the LTM task.*

CWs	*Variables*			
	Concreteness	*Number of letters*	*Frequency*	*CEFR*
Animal	6.0	6	10116	A1
Building	6.3	8	10787	A2
Cake	6.6	4	8740	A1
Jewel	6.2	5	7591	B2
Mouth	6.3	5	10439	A1
Nose	6.5	4	9541	A1
Magazine	6.5	8	11009	A2
School	6.1	6	11695	A1
Street	6.4	6	11240	A1
Salt	6.0	4	9633	A1

***Table 25.** The average rate of concreteness and frequency of CWs in the LTM task.*

Variables	*CWs*
Concreteness	6.29
Frequency	10079

6.2.1.1.2 Participants

The same participants took part in this experiment as in *Experiment 1* and *Experiment 2* (see Section 4.2.1.2*)*.

6.2.1.1.3 Procedure

Before starting the task, the researcher explained the task and provided a tutorial. The participants were given sentences and were asked to write the word that completed the sentences in the form of a cloze test (Ashby-Davis, 1985). The process of completing the sentence with a word is, in fact, a LTM measure (Squire, 1987). Sentences were assessed based on the level of accuracy with which the participants selected the right word to complete the sentences. Speakers relied on their memory to find the right words. Words were presented on A4 size paper, with a text written in black, size 14. The text was written in Times New Roman font (Appendix B). The task took the participants fifteen minutes to complete. No individual feedback was given to participants.

6.2.1.1.4 Scoring and data analysis

Similar to the STM task, scoring the outcomes of this task was conducted by the research and another reviewer to ensure that the results were consistent. This approach was chosen for the LTM task because the scoring procedure involved human interference and judgement. Data were classified based on the type of words. Starting with emotional words, they were further classified as "positive_emotional_words" and "negative_emotional_words". Other groups of words were "non_emotional_words" and "CWs". The process of scoring the outcomes of this task involved the participants' responses according to a set of standards at different levels. The first level of standards was the classification of responses. If a participant came up with a correct AW, he was awarded one point, since the answer was correct. Failure to provide a correct AW in a sentence was scored as wrong with 0 point. The same score was given in case of incomplete/empty responses. Overall, errors in spelling were not penalised. In case the response was synonymous with the required AW, one point was awarded.

Statistically, various approaches were used to test the differences between independent samples using the same testing environment. The selection of any statistical test was based on the homogeneity at the level of selected participants, word selection and the characteristics of the data. It is important in any experimental study to explain the reasons for the selection of tests to increase the power of the results of these tasks. According to Field (2013), these homogeneity concerns push us further to employ parametric tests to examine the difference between AWs. Without considering these reasons in the analysis, it would be impossible to implement these statistical tests. Ideally, these tests should be selected based on their power and trustworthiness in experimental studies. Tests were rationalised based on the assumption that some features of the participants were known (Larson-Hall, 2015). One of the essential rules in using parametric tests is to have data normally distributed with equal sample sizes across groups. As stated in Chapter Four, the advantage of using parametric tests in statistics is the power of the results obtained. There was a possibility of abnormal cases, as this study was conducted on various L2 speakers (Larson-Hall, 2015). It is common to encounter outliers in any experiments because the language being spoken is affected by many factors and individual differences, for example. Therefore, the emergence of these abnormal cases is quite expected, and can be solved statistically. In a large sample size with an equal number of subjects, according to Field (2013), normality tests with only a few abnormal cases do not change the assumption of normality. At this point, there are various factors which affect the normality of the data, and one of them is the ceiling effect. Subsequently, there is a need to run non-parametric tests, depending on the type(s) of the comparison. If the data violate the normality, alternative non-parametric tests will be used instead. In that capacity, mixed-repeated measure ANOVA was the most appropriate parametric test to measure the difference in AWs between beginners and advanced L2 speakers with the same IVs, (Field, 2013; Field & Hole, 2002; Larson-Hall, 2015). This means that both samples were subjected to the same testing conditions. Participants in all the groups were given the same AWs using positive, negative and non-emotional words. Hence, the types of AWs used were repeated measures taken from the selected matched samples. In addition to the previously mentioned requirements in any parametric tests, additional requirements particularly related to this model include:

- Three within-subject factors or IVs should be included, with each one having two categorical levels;
- Data should be continuous; and

- The hypothesis of sphericity was assumed in the populations being sampled. All groups – including the control group – had an equal sample size. It was tested using Mauchly's test in the repeated measures ANOVA which indicates whether the assumption of sphericity is met.[9]

Before analysing the data, the nature of the data reflected different patterns in terms of normality. This feature is essential to determine which statistical test is appropriate for analysing the elicited data. Statistically speaking, the normality of data varied between monolinguals and L2 speakers in this experiment. Starting with the monolinguals, the data showed a ceiling effect, as they were better able to remember various types of AWs. Therefore, it is assumed here that the normality of the data was affected, attributing to the impact of the ceiling effect. This factor led us to analyse the data of monolinguals using non-parametric tests. On the other hand, data was assumed to have a normal distribution across the two groups of L2 speakers with some outliers. Some studies in psycholinguistics "are rarely perfectly normally distributed" (Larson-Hall, 2015, p. 189). Since the number of participants

[9] This test provides values of sphericity which measures the degree of variance in the populations sampled. Mauchly's test ensures that the data fulfil the above-mentioned criteria and validates the sphericity assumption. The violation of sphericity can be solved in many ways. Greenhouse-Geisser and Huynh-Feldt tests are used to adjust the degree of sphericity statistically (Field, 2013). The adjustment of these values can help in assuming the sphericity of data. By deciding to use a mixed repeated ANOVA test to analyse L2 data, the adjustment can be solved. There was an issue of sphericity according to Mauchly's test which was violated (.294). Statistically, the violation was solved by adjusting the degree of freedom in the Greenhouse-Geisser test which was calculated at (F [3.00, 174], $p<.001$). Sphericity could then be assumed under the new adjustments (Field, 2013).

across the three groups was 30, this assists in assuming the normality of the data, even if the normality shape was affected, according to the central limit theorem (Field, 2013). Therefore, dealing with these abnormal cases is quite expected in this field as well as others, especially among L2 speakers. An effort was made in this study to deal with these outliers to run parametric tests. From a statistical standpoint, recruiting a sample with the ($n \geq 30$) may help guarantee the normality of the data, even if the normality shape was affected (Field, 2013). Therefore, differences between these words were further examined to ensure that differences between AWs were statistically significant. The selection of the appropriate test should bear in mind the characteristics of the groups, and whether they received the same AWs. Therefore, the selection of the statistical test here was attributed to:

- Each recruited sample being homogenous in terms of the number of participants in each group, age, gender, right-handedness, education and level of language proficiency;
- The selection of each type of AW being constrained in terms of valence, arousal and number of letters which were assumed to be homogeneous; and finally,
- The independence of each participant not affecting other participants.

The statistical approach included both descriptive and inferential statistics. For the effect of sample size, the analysis includes partial eta squared (η_p^2) as an outcome of repeated measures ANOVA test.

6.2.1.1.5 Results

The results for AWs are presented sequentially, starting with positive, negative and ending with non-emotional words compared to CWs. In terms of the IVs among L2 groups, there were two IVs in this experiment word type and the level of language proficiency. For the monolingual group, word type was the only IV. The number of remembered words was the DV of the experiment for all the groups.

Starting with identifying the means of remembering AWs, it is worth mentioning that reporting the statistics happened at two stages. The first stage aimed to report the means of positive and negative emotional words, and examined whether there was a significant difference between these types of words. If so, the amount of remembered words was converted into proportions, since the number of positive and negative emotional words was ten (five for each). For L2 speakers, the results showed that positive emotional words were remembered better (M=2.06, SD=1.59) if compared to negative emotional words (M=1.36, SD=1.32) among L2 beginners. The same pattern of memory remembrance emerged among advanced speakers in remembering positive emotional words better (M=3.26, SD=1.33) compared to negative emotional words (M=2.73, SD=1.25). To make sure these differences were significantly different, inferential statistics were conducted between positive and emotional words among L2 speakers. Since the comparison was between these two types of words, a repeated ANOVA was used to examine whether the differences between these words were significant. The assumption of sphericity as stated in Mauchly's test was met by choosing this test because it was run between two levels of comparison between positive and negative emotional words. The results indicated that there was a difference between positive and negative emotional words, $F(1, 58) = 14.88$, $p<.001$, partial $\eta p2 = .204$. Language proficiency between beginners and advanced L2 speakers reflected a significant main effect ($p<.001$) which, therefore, affected the recallability of emotional words. Even the results of partial eta squared showed that the impact of sample size was insignificant between positive and negative emotional words with no effect ($p=204$).

Reaching these results led to compute all the values into proportions to avoid any potential effect of the sample size of word in non-emotional words (n=10) and CWs (n=10) compared to positive (n=5) and negative emotional words (n=5). It would be incorrect to run a parametric test without having an equal number of words for each type of word. In STM task, this action was impossible since these words were treated as one type of words compared to non-emotional and CWs. At the descriptive level, CWs among advanced L2 speakers were remembered better in this task (M=.77, SD=.15), similar to non-emotional words (M=.73, SD=.16). The performance of beginners showed an unusual pattern for non-emotional words (M=.33, SD=.22) in comparison to both negative emotional words (M=.27, SD=.26). Conversely, advanced L2 speakers showed a high level of remembrance of positive (M=.68, SD=.20) and negative emotional words (M=.54, SD=.20) compared to L2 beginners. CWs were remembered the most compared to all three types of AWs. Similar to STM, CWs were remembered at a superior level by both beginners and advanced speakers (Table 26).

Table 26. *Means of remembering various types of AWs compared to CWs in the LTM task among L2 groups and monolinguals(n=90).*

Word Type	*Group(s)*	*M*	*SD*	*Mid*
Positive Emotion Words	Beginners	.4133	.31919	.400
	Advanced	.6867	.20800	.700
	Monolingual	.96	.096	1.00
Negative Emotion Words	Beginners	.2733	.26514	.200
	Advanced	.5467	.20297	.600
	Monolingual	.833	.149	.800
Non-Emotional Words	Beginners	.3333	.22642	.300
	Advanced	.7300	.16640	.700
	Monolinguals	.903	.076	.900
CWs	Beginners	.6033	.23413	.600
	Advanced	.7700	.15347	.75
	Monolinguals	.906	.094	.900

In this analysis, there were two directions of difference. The first was between the types of AWs, and the second was between AWs and the level of language proficiency of L2 speakers. At first, this model showed that there was a difference between AWs, $F(3, 174) = 24.24$, $p < .001$, partial $\eta p2 = .295$. The results of these pairwise comparisons showed a multitude of differences between positive, negative and non-emotional words ($p<.001$). Specifically, remembering positive emotional words was significantly different from remembering negative emotional words ($p=.002$), but not in comparison to non-emotional words ($p=.994$). For negative emotional words, the results showed a difference between negative emotional words and non-emotional words ($p=.003$). All three types of AWs showed significant differences compared to CWs ($p<.001$). Regarding the level of language proficiency, a comparison was made between beginner and advanced L2 speakers, and the result demonstrated a significant difference ($p<.001$). Both types of emotional words demonstrated significant differences. Starting with the positive emotional words, the outcomes of the pairwise comparison indicated that there was a significant difference between beginners ($p<.001$) and advanced L2 speakers ($p<.001$). For negative emotional words, there was a significant difference between beginners ($p<.001$) and advanced L2 speakers ($p<.001$). These differences were also observed in non-emotional words. There was a significant difference in remembering these words among beginners ($p<.001$) and advanced ($p<.001$).

Even with CWs, the difference continued to be significant ($p<.002$) between these two levels of language speakers.

A more advanced comparison was also conducted that aimed to further examine the interaction between the level of language proficiency and the recallability of AWs. The outcomes of the same statistical model measured whether there was an interaction between recallability of AWs and level of language proficiency. The results showed that there was an interaction between the two IVs, $F\ (3.174) = 4.176$, $p =007$, partial $\eta p2 = .0.06$. More patterns of interaction appeared in the remembrance of these words when further analysis was conducted on both the beginners and advanced L2 speakers.

Overall, different patterns of memory in LTM emerged compared to STM, suggesting different patterns of interaction between the remembrance of AWs and the level of L2 speakers. Therefore, these patterns can be grouped into: dominant and secondary. The first pattern emerged in the differences between emotional (positive ↔ negative) words among beginners and advanced L2 speakers. Both types of emotional words showed a consistent pattern of memory with an interaction with the level of language proficiency. The second pattern was found in the non-emotional words. Clearly, non-emotional words showed various patterns of LTM, depending on the level of language proficiency. Among L2 beginners, these words did not differ significantly from negative emotional words ($p=.205$), but they showed a difference compared to positive emotional words (p=.082). Among advanced speakers, non-emotional words showed the opposite, being different compared to negative emotional words (p<.001), but insignificant compared to positive emotional words (p=.342). All three types of AWs showed differences compared to the control words, except for non-emotional words among advanced L2 speakers (p=.297). Patterns of interactions in LTM can be summarised briefly in a different way here between:

- Positive and negative emotional words among beginners and advanced L2 speakers;
- Positive emotional and non-emotional words among beginners;
- Negative emotional and non-emotional words; and
- CWs and all AWs among beginners and advanced L2 speakers (Table 27).

Table 27. *Repeated measure ANOVA between AWs and the level of L2 proficiency (n=60).*

L2 Groups	*The direction of the Interaction*		*P*	*Sig.*
Beginners (n=30)	Positive emotional words	Negative emotional words	p=.008	***
		Non-emotional words	p=.082	***
		CWs	p<.001	***
	Negative emotional words	Non-emotional words	p=.205	
		CWs	p<.001	***
	Non-emotional words	CWs	p<.001	***

L2 Groups	*The direction of the Interaction*		*P*	*Sig.*
Advanced (n=30)	Positive emotional words	Negative emotional words	p=.008	***
		Non-emotional words	p=.342	
		CWs	p=.091	***
	Negative emotional words	Non-emotional words	p<.001	***
		CWs	p=.001	***
	Non-emotional words	CWs	p=.297	

Significant level of p value refers to either: *=.001, .01=**, or .05= *.*

For the monolingual speakers, a further comparison was made to determine whether the differences between emotional and non-emotional words were statistically different. Similar to L2 speakers, data was converted to deal with differences in the word list in a similar way. Due to the abnormality of the data, the selection of a non-parametric test was attributed to the ceiling effect of the data. Monolingual speakers were able to show a superior level of memory performance compared to L2 speakers. This superiority was observed in remembering all three types of AWs alongside CWs. Statistically speaking, Friedman's test χ^2 was employed to examine these differences, and the results of this test showed that there was a significant difference between these words χ^2 (3, n=30) = 16.992 p<.001. The median indicated that both positive emotional words were the highest in remembrance (Mdn= 1.00). Both non-emotional words and CWs had the same level result (Mdn= .900), whereas negative emotional words were remembered the least (Mdn = .80). The results of the pairwise comparison suggested different patterns. Between positive and negative emotional words, there was a significant difference between these two types of words ($p=.002$). This was the only significant difference among monolinguals. No significant difference was observed between positive and non-emotional words on the one hand ($p=.129$), and between negative and non-emotional words on the other ($p=1$). Even for CWs, no significant difference was observed between CWs, emotional and non-emotional words.

After determining the results for remembering various types of AWs among L1 and L2 speakers, it was important to make another dimension of comparison at the word level. This comparison aimed to compare the pattern of remembrance across groups. Since the nature of the data differs between monolinguals and L2 speakers, a different type of statistical test was employed. To explain, the data of L2 speakers were analysed using a parametric test, whereas the monolinguals' data were analysed by a non-parametric test. In this case, comparing the word types employed a non-parametric test since the data violated the assumption of normality for monolinguals. Therefore, the three types of AWs were examined using a Kruskal-Wallis test across groups. Starting with positive emotional words, the results showed that there was a significant difference in remembering these words among the three groups $\chi^2 (2, 90) = 48.00$, $p<.001$. As confirmed by pairwise comparison, there was a significant difference between beginners and advanced L2 speakers in remembering these words ($p=.026$). Compared to monolingual speakers, the results demonstrated a significant difference between beginners and advanced speakers, ($p<.001$) and ($p<.001$) respectively. Similarly, the recallability of negative emotional words showed significant differences across the three groups χ^2 (2, 90) = 48.53, $p<.001$. The pairwise comparison result showed a significant difference in remembering these words between L2 beginners and advanced speakers ($p=.007$). The same significant difference was found between beginners and advanced speakers on the one hand ($p<.001$), and between advanced speakers and monolinguals on the other ($p<.001$). For non-emotional words, the same patterns of remembrance emerged across the three types, even though this type of word is semantically different. The same statistical test was employed, proving the significant difference in remembering these words χ^2 (2, 90) = 57.69, $p<.001$. The results of the pairwise comparison demonstrated that there was a significant difference in remembering these words among beginners and advanced L2 speakers ($p<.001$). Compared to monolinguals, there was a significant difference between beginners and advanced speakers ($p<.001$) for both groups. Even for CWs, differences in remembering these words were significant across the three groups χ^2 (2, 90) = 31.84, $p<.001$. As pairwise comparison showed, there was a significant difference in remembering these words between beginners and advanced speakers ($p=.032$). A significant difference remained between both monolinguals and advanced speakers ($p=.006$) and monolinguals and beginners ($p<.001$).

Important findings, therefore, emerged from this test which indicated that emotional words were remembered better than non-emotional words and CWs were remembered better than non-emotional words. Non-emotional words reflected unsystematised patterns compared to emotional words and CWs across groups.

6.2.1.1.6 Observations in the experiment

Whilst there were many initial justifications for the selected design, important insights emerged from using the cloze test (particularly in terms of memory patterns, as shown above). It became evident following the experiment and its analysis which an additional experiment was necessary to deliver more reliable insights into the recallability of AWs in the LTM. At the design level, it was apparent, for instance, that the design did not prioritise the importance of the learning phase before running the experiment which may raise concerns around the participants' linguistic ability. At the participants' level, it was possible that some of these participants would now know the words. Furthermore, there was another concern that recruiting participants with a high level of language proficiency was not well-justified because their linguistic ability was high, and this may cause a ceiling effect. The same applies to monolingual speakers of English who, for example, would have no problem with addressing all items in the cloze test, whereas beginners would face difficulties in remembering the expected words. The use of the cloze test in English with the selected participants did not allow for a sufficient comparison between monolinguals and L2 speakers. At the word level, where some English words are used as both nouns and adjectives, it was difficult to determine the distinction in the retrieval process between some of the words in the test. Even with the effort to create a context where only AWs were needed, there were concerns about the interchangeability in remembering abstract nouns only. Another concern was related to the classification of AWs provided by the participants. It was apparent that the task did not allow control of words, as some participants may provide synonyms. Although this behaviour was expected before running the experiment, it was not clear if the number of remembered words was from the previously selected list in the experiment or not. Thus, the new experiment aimed to attend these limitations, and ensure that any gaps in understanding the mechanisms of remembering AWs in LTM were addressed.

6.2.2 The use of multiple-choice and translation tests to measure the recallability of abstract words in the long-term memory

As touched on earlier in the thesis, previous studies on LTM experiments were based on measuring the recallability of AWs using two approaches: psychological approach (Paivio & Csapo, 1969) and one based on L2 learning studies (Bogaards, 2001). As stated in Chapter Three, the former focused on presenting words to participants, asking them to remember them after a week. Procedurally, words were retrieved without providing any cues, so the task was in the form of free memory task. The latter, which influenced Experiment 4 in the current study, was adapted from L2 learning studies (Bogaards, 2001). It provides another perspective on measuring the recallability of words in the LTM compared to multiword expressions. The recallability of words is tested by using multiple-choice and translation tests. Again, the selection of these two tests was rationalised, emphasising the role of context in the retention and retrieval stages. This is realised in the way these words are presented to the participants in the learning and examination stages.

Implementing this design in measuring the recallability of AWs was considered appropriate, especially in relation to how words were retrieved. At this point, there were two ways in implementing this design. At first, participants were provided with choices, and asked to select one of the words which described/defined the intended word. In this case, the design controls the mechanism of selecting the right choice. Only one answer is acceptable. Secondly, the design linked the recallability of words to the participant's native language by providing another test which is about translation. This is based on the fact that storing words, either newly or previously learned words, is stored in the LTM with their associations. Part of these associations is the translation of these words. Another consideration in this design was the characteristics of the sample. Following Bogaards'study (2001), it was decided that choosing monolinguals and exposing them to a language was the most appropriate here to examine how AWs intend to be stored in the LTM. This would involve introducing different types of AWs and comparing them with those of L2 speakers with a modest level of the language. The use of these two tests has some advantages and disadvantages.

Starting with the merits, the adopted design considered the importance of the learning phase in measuring the recallability of AWs. In LTM tasks, the focus is on measuring whether or not these words are retained in the memory for a long period of time. It is hard to have LTM task without learning and examination phases. These two phases in measuring LTM were missing from Experiment 3. In fact, including these two phases in this design applied to the characteristics of measuring the recallability of words in the LTM.

On the other hand, this design has some shortcomings. Contextual factors in the learning and examination phases, like lexical richness, may raise concerns. These factors may affect the participants' performance in remembering the intended word(s). The degree of difficulty, as well as, the context of the topics, also play crucial roles in the way speakers remember words. These factors are centralised around the word choice and depend on the characteristics of the words which need to be measured. The topic of each test item is another dimension which could affect memory performance. As stated in Experiment 3, a topic that is highly sophisticated in physics or neuroimaging makes the speakers' job difficult in remembering the required word even if the intended word is easy to learn (Daller *et al.,* 2007). Another example of the effect of context is the length of sentences, and whether the text in explaining the intended word should be long or short. What makes the difference is the number of words given to explain the meaning of the word. If the meaning of a word has been explained in one or two sentences, it is different from if the meaning has been explained in a whole paragraph. To characterise this difference, there are three interrelated issues: the level of language, validity and reliability of the proposed context. Using highly sophisticated words in a context which describes the meaning of a particular word makes the job of retrieving a word more challenging. Instead of thinking about the intended word, efforts are made to think about the meaning of these difficult words. This issue becomes more challenging when measuring LTM amongst speakers who speak more than one language. Despite the potential limitations of the multiple choice and translation test, this was considered to be more effective than the cloze test in achieving the data required. It was important, however, to consider these issues in the design of Experiment 4 (see below).

6.2.2.1 Experiment 4

In this experiment, the IV was the type of words, excluding the level of language proficiency. The DV, on the other hand, was the number of remembered words. This experiment aimed to answer the same research questions raised as Experiment 3 which were listed as follows:

1. Similar to STM task, how do speakers of two languages with different levels of linguistic abilities remember emotional and non-emotional words in LTM? How they are different compared to monolinguals?
2. Compared to CWs, what are the memory patterns of remembering emotional and non-emotional words in LTM?

Similar to the STM task, there were also two hypotheses to be examined here:

1. Emotional words are better remembered compared to non-emotional words among L2 speakers in LTM; and
2. CWs are remembered better than emotional and non-emotional words LTM.

6.2.2.1.1 Design and materials

The same list of words was used in this experiment as in Experiment 3, although the design of this task was different (Table 19, 20, 22, & 24). The context being used was taken from Cambridge Dictionary of English. These items were translated into Arabic since they would be presented in Arabic. The translation being reviewed by native speakers of Arabic to ensure it was clear and understandable (Appendix C). The theme or the topic for the test items describes daily life topics rather than highly sophisticated topics since the main focus here was to measure the recallability of various types of AWs.

6.2.2.1.2 Participants

Participants were divided into two groups. The first group included 10 male L2 speakers with English as their L2 and Arabic as their L1. Particular characteristics were sought in relation to English language proficiency, gender, age and physical abilities. Starting with the level of language proficiency, this task aimed to recruit beginners whose IELTS scores ranged between 4 and 5. According to the British Council, participants between these two scores are classified as beginners. In terms of gender, the selection of participants was limited to males due to cultural considerations. While it may have been possible for a male researcher to recruit L2 females in the UK, this would have been more difficult in the Saudi context. Participants' age range was between 20-30. This selection was similar to other experiments in the thesis to eliminate the effect of age on the elicited data. For the physical abilities, participants did not have any known physical, psychological or neurological issues, or any visual or hearing impairments. The second group included 10 Saudi male monolingual speakers of Arabic who had lived in Saudi Arabia all their lives. Except for language proficiency, all of the above characteristics also applied to this group.

6.2.2.1.3 Procedure

The design of the LTM task followed Bogaards' study (2001) in measuring the retention of words. The number of words in this experiment was 30; 10 for each type of AW. Measuring the recallability of these words occurred at two phases:

learning and examining. The learning phase took place first. Participants were taught the AWs alongside CWs, including how to pronounce the AWs, as well as their meaning and usage. Providing the translation of each word was also made at this phase to remove any confusion with other words. Further follow-up questions at the end of this phase were used to ensure participants fully understood the meaning of the AWs. At the same time, they were free to ask any questions, regarding the usage of these words. During this phase, participants were asked to write down the Arabic translation of each word. The duration of this phase was one hour. Next was the examination phase. This phase happened after one week. Participants were asked to return for another session without telling them they would be asked about the words presented before. They were not allowed to go back to the list of words presented and refresh their memory. Two types of tests were used to examine the recallability of the words: multiple-choice and translation tests, according to Bogaards' design (2001). In the multiple-choice test, participants were given an explanation for each word in Arabic, followed by three alternatives in English. It was the job of the participants to choose the word that described the meaning correctly. The second type of test was a paper-based translation test (see Appendix C). Distractors were provided in this experiment from the words which had been used during the learning phase. These distractors varied between emotional, non-emotional and CWs. Participants were given words in

English, and asked to provide the Arabic translation. The duration of this phase was 20 minutes.

6.2.2.1.4 Ethical considerations

While this experiment was conducted later in the study, it was still important to address ethical issues. A request for ethical approval was, therefore, submitted to the ethics committee in the Department of Language and Linguistic Science at the University of York. The request included all the requirements for collecting data, including issues of consent, payment, safety, confidentiality and so forth, and was approved prior to data collection.

6.2.2.1.5 Difficulties in data collection

Since this experiment was conducted later in the thesis, there were a few challenges in collecting data. These challenges were categorised into challenges related to the sample being recruited and those related to the time. Starting with the sample, some challenges were faced during data collection in finding enough participants in a limited time. Since deciding to run another LTM experiment was made before the thesis submission, it was hard to find enough participants in Saudi Arabia whose characteristics suited the criteria of the study, especially given the timing of the study (during a semester break) and the additional time pressure. There was an option to travel to other cities in Saudi Arabia, however, this required a budget to commute across the country.

6.2.2.1.6 Scoring and data analysis

The participants’ responses were classified according to clear criteria in this task. Correct choices and acceptable translations were given one point and wrong answers or no answer received zero. Spelling mistakes were not penalised. Coding data for the analysis followed the same method as that adopted in Experiment 3. The nature of data in this experiment showed lack of normality as the histogram confirmed the assumption of abnormality (Figure 8).

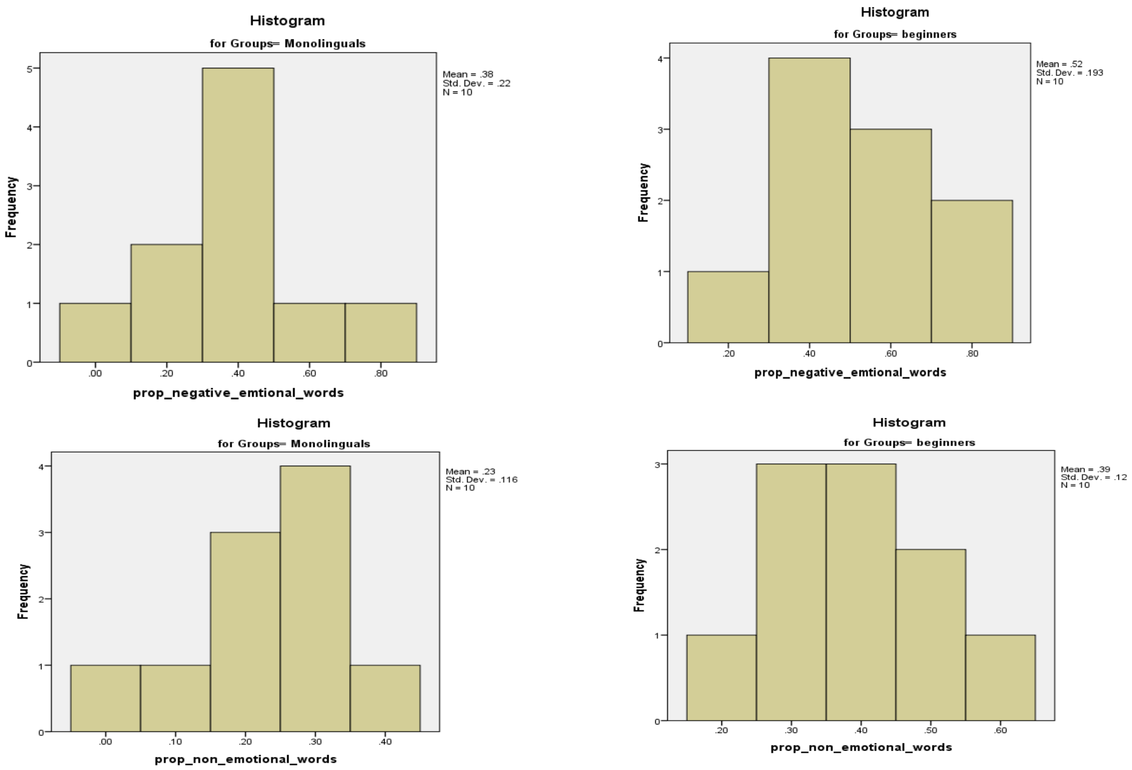

Figure 8*. Characteristics of the data in the LTM task.*

6.2.2.1.7 Results

After running this experiment on monolinguals and L2 speakers, different patterns emerged in examining the recallability of emotional and non-emotional words. Inferential statistics were used to determine whether differences in remembering these words across groups were statistically different or not. Before reporting the results of memory patterns of AWs; however, it is essential first to determine if there was a difference in recall within emotional words. This step is important since the number of positive and negative emotional words was not even (n=5). At this point, there were two possibilities. If the difference in remembrance was statistically different, each type of emotional words would be separated in comparison. If there was no difference in remembrance, emotional words would be treated as one group.

The descriptive statistics indicated that positive emotional words were remembered better among beginners (M=3.7 SD=.67) than monolinguals (M=2.3 SD=.94). Such superiority in remembering emotional words was also found in negative emotional words. Beginners were able to remember negative emotional words better (M=2.6 SD= .96) than monolinguals (M=1.9 SD= 1.1). In both groups, the number of participants plays a role in selecting the type of statistical test. According to Field (2013), selecting a sample with less than 30 subjects may not lead to collect data normally distributed. Since the number of each group was below 30, a decision was made to run a non-parametric test to determine if the difference between positive and negative emotional words was statistically different. As the comparison was between the two groups in examining positive emotional words, the results of Mann-Whitney U test indicated that there was a significant difference in remembering these words between these two groups The result of this test indicated that there was a significant difference between beginners (Mid = .80) and monolinguals (Mid = .46) in remembering positive emotional words, U= 12.50, p=.003. For negative emotional words, there was no significant difference in remembering among beginners (Mid=.50) compared to monolinguals (Mid=.38), U=31.50, p=.165. Another dimension in examining the difference between these words was within each group. The result of Wilcoxon test (W+) showed that there was a significant difference in remembering positive emotional words (Mdn=.80) compared to negative emotional words (Mdn=.40) among beginners, z = -2.23, p=026. There was no significant difference between positive (Mdn= .50) and negative emotional words (.40) among monolinguals, z = -945, p=.34. Having demonstrated these findings between positive and negative emotional words, a decision was made to convert the amount of remembered words into a proportion. This is based on a significant difference in remembering these two types of emotional words between the two groups. On the other hand, the number of positive and negative words in this experiment compared to non-emotional and CWs solidified this decision too. After converting these results into a proportion, the results indicated that positive emotional words were remembered better by beginners (M=.74 SD=.13) compared to monolinguals (M=.46 SD= .18). This pattern continued between the two groups in remembering negative emotional words. The results showed that negative emotional words were remembered better (M=.52 SD=.19) than monolinguals (M=.38 SD= .22). Even with non-emotional words, beginners continued to have the same memory pattern in remembering these words better (M=.39 SD=.11) compared to monolinguals (M=.23 SD= .11). For CWs, monolinguals had a higher level of remembrance compared to both types of emotional and non-emotional words (M=.54 SD=.13), but beginners were still able to remember CWs better (M=.74 SD=.13) (Table 28).

Table 28. *Means of remembering various types of AWs compared to CWs in the LTM among L2 groups and monolinguals(n=20).*

Word Type	*Group(s)*	*M*	*SD*	*Mid*
Positive Emotion Words	Beginners	.74	.13	.80
	Monolingual	.46	.18	.40
Negative Emotion Words	Beginners	.52	.19	.50
	Monolingual	.38	.22	.40
Non-Emotional Words	Beginners	.39	.11	.40
	Monolinguals	.23	.11	.25
CWs	Beginners	.74	.13	.80
	Monolinguals	.54	.13	.60

To confirm if these differences were statistically significant or not, inferential statistics were used and listed here according to groups and word type. Starting with groups, the results of Friedman's test χ^2 demonstrated that there was a significant difference in remembering AWs among beginners. The results of χ^2 showed differences between the three types of AWs. The results revealed a significant difference in scaling these words, χ^2 (3, n=10) =21.90,

$p<.001$. The median indicated that both positive emotional words and CWs were the highest in emotionality (Mdn = .80) compared to negative emotional words (Mdn = .50). Non-emotional words were rated the lowest in this scale (Mdn = .40). According to the outcomes of the pairwise comparison, there were different patterns of differences between these words. To specify, there were significant differences between positive emotional and non-emotional words ($p=.003$) on the one hand, and between non-emotional words and CWs on the other ($p=.001$). The results of the post hoc test did not report any significant difference between both types of emotional words and CWs or between both types of emotional words. Also, there was no significant difference between negative emotional words and non-emotional words.

As for monolinguals, the results of Friedman's test (χ^2) test showed a significant difference in scaling these words, χ^2 (3, n=10) =16.159, $p<.001$. The median indicated that CWs were the highest in remembrance (Mdn = .60) compared to all AWs. Non-emotional words were rated the lowest in this scale (Mdn = .25). As for positive emotional words and negative emotional words, both types were remembered the same (Mdn=.40). The outcomes of the pairwise comparison indicated that the only significant difference in this group was found between non-emotional words and CWs ($p=.001$). As the test examined other types of words, there was no significant difference between emotional words ($p=.363$). A lack of significant difference was also observed in both types of emotional words and non-emotional words. To specify, there was a difference between positive emotional words and non-emotional words ($p=.009$). The same applied to negative emotional words and non-emotional words ($p=.083$), and between both types of emotional words and CWs. The results showed that there was no significant difference between positive emotional words and CWs ($p=.299$), and between negative emotional words and CWs ($p=.057$).

Another dimension of the comparison here is at the word level between groups. The Mann-Whitney U test was employed since we are comparing the recallability of each word type across the two groups. The result of this test indicated that there was a significant difference between beginners (Mid = .80) and monolinguals (Mid = .40) in remembering positive emotional words, $U= 12.50$, $p=.003$. This pattern was also observed in remembering non-emotional between beginners (Mid =.40) and monolinguals (Mid =.25), $U = 17.00$, $p=.011$. Even in remembering CWs, the same pattern of significant difference was also in CWs between beginners (Mid =.80) and monolinguals (Mid = .60), $U = 15.50$, $p=.007$. On there other hand, there was a lack of difference in remembering negative emotional words among the same groups respectively (Mid= .39, Mid= .23), $U=31.50$, $p=.165$.

6.3 Summary

The analyses of the LTM experiment 4 across the two groups showed that the recallability of emotional and non-emotional words in LTM varied. Some of the differences that emerged were expected, such as positive emotional words being remembered better than non-emotional words. Others showed unexpected memory patterns. At this stage, interpreting the emergence of these patterns in this chapter as well as in the previous two chapters represents the essence of the next chapter. It will provide a detailed discussion of the patterns of AWs, relating to the distribution of emotionality and storing of emotional and non-emotional words in STM and LTM. As for the patterns emerged in experiment 3, the discussion in the following chapters will not address these patterns since the design had a few concerns which could lead to inaccurate interpretations.

Chapter 7

Discussion

7.1 Introduction

The findings, thus far, indicate that various types of AWs are not patterned in the same way in terms of emotional intensity and retaining words in the memory. It has been shown that similarities and differences exist between monolinguals and L2 speakers. The classification of these words alongide the methodology employed for this study helped to demonstrate the recallability of various types of AWs and the reason(s) which affect the emergence of these patterns. This chapter begins by discussing the results of the emotionality rating in a separate section. The section discusses the emergence of emotionality patterns of AWs among two types of the lexicon, explaining the justification behind the appearance of these patterns. After that, the outcomes of the STM and LTM experiments will be discussed together since both of them measure the recallability of AWs but from different perspectives. The discussion highlights the impact of several factors, affecting the emergence of various patterns. This is followed by a separate section on the impact of emotionality on the recallability of various types of AWs. The section examines the relationship between emotionality and memory patterns of AWs by revisiting the current hypotheses. The chapter concludes by addressing the study's limitations, contributions and implications for the field.

7.2 Emotional Intensity of Various Types of Abstract Words

In this section, there will be a discussion on the distribution of emotionality for various types of AWs. At the micro level, discussing the findings is organised in this chapter based on providing a summary of the existing literature about the emotionality of AWs with a description of what has been missed. The summary includes the findings of the thesis, including similarities and differences between the current research and the existing literature. After that, interpreting the current research findings and linking them to the emotionality model by Kousta and Vigliocco are provided. The discussion concludes by addressing the mini-research questions and (dis)proves hypotheses raised earlier in Chapter Four.

7.2.1 Emotional intensity varies according to word type

As touched on in Chapter Three, whilst there has been a range of literature which has explored the emotionality of AWs (Siakaluk *et al*., 2014, 2016; Anooshian & Hertel, 1994; Ferré *et al.* 2010; Ferré *et al.*, 2018), there has been relatively little attention given to measuring the emotionality of these words among speakers who can speak another language as L2 but not bilinguals. As one feature of studying this factor is the strength of emotional intensity. There has been some debate surrounding whether the impact of emotionality is stronger among L2 speakers. Research on bilinguals suggests that they have one framework in rating emotionality (Ayçiçeği & Harris, 2004; Ferré *et al*., 2010, 2013), whereas research by Ferré *et al.* (2018) argues that emotionality in the speakers' native language should be stronger since it is the language used with their families from early childhood. Determining the level of emotionality among L2 speakers had not previously been achieved, making it an important gap in knowledge and something addressed in the current study. It was found that there were significant differences between emotional and non-emotional words among L2 speakers and monolinguals. The results related to L2 speakers demonstrated that emotional and non-emotional words had been scaled similarly to monolinguals. Within emotional words, there was no significant difference between the advanced L2 group and monolinguals. However, differences emerged between the three groups in rating non-emotional words.

Taken together, this suggests that rating the emotionality of AWs has two patterns, one with emotional words and one with non-emotional words. These findings showed some consensus with the literature, regarding rating emotional words higher than non-emotional words and a lack of differences within emotional words (Siakaluk *et al*., 2014, 2016). This finding was also observed among bilinguals regardless of whether they acquired the L2 early or late (Anooshian & Hertel, 1994). There was also some consensus with the literature in relation to the lack of difference between positive and negative emotional words (Ferré *et al.,* 2010; Siakaluk *et al*., 2016). As for the strength of emotionality, the outcome of the current research did not agree with Ferré *et al.*'s claim (2018), as the current study suggested that the strength of emotionality showed no difference across groups. Therefore, this research contributed a significant finding compared to previous studies.

There was a speculation prior to conducting the experiment by the researcher that emotional intensity would take two patterns. The first pattern indicates that there will be an increase in strength when a speaker acquires another language. This is justified because L2 speakers have more emotional associations from the language they acquire in addition to their native language. On the other hand, another pattern characterises emotionality among monolinguals, as stronger since it is acquired earlier, and it is retained in the memory during the early childhood of the speaker. The findings from the current study contradicted with these two patterns. The current study clarifies the similarity between monolinguals and L2 speakers in rating emotional words together with the strength of emotionality. At this point, it can be argued that the semantic characteristics of emotional words had an impact on establishing/linking the level of emotionality more to emotional words. Some AWs were expected to possess a particular quality, varying between emotional and non-emotional words. A word, like *joke,* as a positive emotional word, is linked to funny stories, beautiful experiences or happy memories. Compare the word *joke* to a word, like *concept,* as a non-emotional word, the semantic characteristics of emotional words allow the emotional associations to be linked to emotional words more than non-emotional words. The semantic characteristics of non-emotional words, on the other hand, do not refer to emotion. As a result of this distinction, we can see that even L2 speakers had the same pattern in rating emotional words higher than non-emotional words. The similarity in these patterns is important to highlight, as the current research drew on a different sample, focusing on L2 speakers with different levels of linguistic abilities rather than true bilinguals. At the same time, this suggests that the speaker's native language directs the establishment of emotionality. This is reflected in rating emotional words and in the strength of emotionality. Interestingly, the strength of emotionality reflects an unexpected pattern which emphasised the role of the speaker's native language in directing the level of emotional intensity, even among L2 speakers, with no difference between monolinguals and L2 speakers. Taken together, the semantic characteristics of emotional words and the speaker's native language indicate that the lexicon of L2 speakers establishes emotionality based on their L1 since the semantic meaning of emotional words in their L1 allows them to acquire emotionality similar to monolinguals.

Another pattern emerged in the current study with emotional words was the lack of difference between positive and negative emotional words. The selection of positive and emotional words ranged between fixed limits, stated at the methodological stage of the study. These restrictions led to observe no difference between positive and negative words in rating the level of emotional intensity. Theoretically, if the selection of these words did not consider controlling the degree of valence or arousal, the difference between positive and negative emotional words would be clear.

From a broader perspective, the findings emerged from the current study can be interpreted by looking at the work of Kousta and Vigliocco (2010) on the representation of emotional words. The model presented in Chapter Two showed the same results in rating emotional words higher than non-emotional words among monolinguals. The similarity in scaling emotional words was expected among monolinguals since monolinguals stand in this study as a control group. By doing this, we can trust the findings of the scale since it came in line with the outcome of Kousta and Vigliocco's model. As for the L2 speakers, the model can be extended to characterise the emotionality of emotional words since it was studied on a different sample. One can argue here that the distribution of emotionality varies at the word level among L2 speakers. Therefore, the current study's findings, which emphasised the significance of word type in rating emotionality, contribute to Kousta and Vigliocco's model (2010) model. This is a significant contribution, as one of the limitations of Kousta and Vigliocco's model, which has been highlighted in Chapter Two, is the extendibility of the model to L2 speakers.

These insights on rating emotional words have helped to address the first research hypothesis (1) stated in Chapter Four. It was hypothesised that emotionality varies between monolinguals and L2 speakers. Previously discussed patterns about the variations between emotional and non-emotional words in the current study gave us the confidence to accept this research hypothesis. Originally, the hypothesis was directional since it assumed the variation of emotionality within various types of AWs. Put differently, the hypothesis assumed that emotionality varied between various types of AWs on the one hand, and between L2 speakers and monolinguals on the other. However, the current study found no difference between monolinguals and L2 speakers. Therefore, this part of the hypothesis could not be accepted based on the study's finding. From this, we can look at the first mini research question of the scale (1) which questioned how L2 speakers rate various types of AWs compared to monolinguals. The outcomes of the scale suggested that variations in rating the emotionality of AWs emerged at word level rather than the kind of lexicon. On the other hand, the insights of the strength of emotionality, which represents the second mini research question of the scale rating (2), came to also confirm that the strength among L2 speakers is directed by the semantic characteristics and the speakers' L1. The scale findings indicate that emotionality, as a non-linguistic factor is not affected by the speaker's linguistic ability. It is a psychological factor with no influence on the speaker's language proficiency in increasing the strength of emotionality. This was a knowledge gap in the literature, questioning the establishment of emotionality among L2 speakers with different linguistic abilities. Based on these insights, it is hard to prove the second research hypothesis (2) stated in Chapter Four which assumed that the strength of the emotionality of L2 speakers is higher than monolinguals.

7.2.2 Non-emotional words patterned differently

The existing literature on non-emotional words in Chapter Two has shown that the representation of these words is affected by many factors, including the degree of concreteness and emotionality (Altarriba *et al.,* 1999; Kousta *et al.,* 2009, 2011; Vigliocco *et al.*, 2010, 2013). Previous work agreed that non-emotional words have proved to have low or no concreteness and emotionality compared to emotional words. The semantic features of these words lead to scale these words low against these norms, even among bilinguals (Ayçiçeği & Harris, 2004; Ferré *et al*., 2010, 2013). Among L2 speakers, the current study showed a significant difference in rating non-emotional words compared to emotional words. The direction of the difference in the current study was not systematic in the sense that amongst L2 speakers, there was no difference in rating these words. The difference emerged between monolinguals and beginners only. These findings contribute to the patterns of these words across groups. Some of these findings showed a similarity with the existing literature in rating non-emotional words lower in terms of emotional intensity compared with emotional words (Ayçiçeği & Harris, 2004; Ferré *et al*., 2010, 2013). Where the current research differs, rating the emotional intensity of non-emotional words varied within L2 speakers. Remarkably, the current study displayed that those speakers with a modest level of L2 rated these words higher than monolinguals. At the same time, this level of intensity did not show any significant difference between monolinguals and advanced L2 speakers. This was unexpected, as it was anticipated that advanced L2 speakers had higher levels of emotional intensity in the current scale.

What can be learned from these findings is that the semantic characteristics of non-emotional words are the reason behind rating these words low. Semantically, some of these words refer to thought, faith and logic, and therefore, they are termed logical words (Khokhlova, 2014). Others are considered to be emotional descriptors, as they are linked to emotional associations. To give examples, a word, like *memory,* as a non-emotional word, reflects a positive emotion for someone with lovely childhood memories, and therefore it is rated higher than a word, like *concept,* even though both words are classified as non-emotional words. Taken together, it seems that it was hard to find a link between emotionality and words, like *time* or *total*.

This non-systematicity is attributed mainly to the semantic meaning of these words, leading to the establishment of various associations with these words. From this, one can also argue that the behaviours of these words are affected by other factors. One of these factors is the level of word difficulty, as some L2 speakers may encounter difficulties in understanding the meaning of some words. Consequently, rating the emotionality of these words is affected. In Chapters Four, Five and Six, there was a clear statement about selecting emotional and non-emotional words based on thematic reasoning. As a result, the selection of AWs was based on choosing words with various levels of word difficulty according to CEFR[10]. It was difficult to control the selection of these words between fixed ranges to eliminate their effect on measuring emotionality of AWs. In the current research, most of the positive emotional words varied between B1 and B2, according to CEFR. Only *acceptance* and *glory* were

[10] Words in this framework are divided into three levels: A, B, and C, where C represents words with the highest difficulty level, and A represents the lowest level. Words in the B level are classified in the middle in terms of the level of difficulty.

classified as C1. For negative emotional words, only the word *hate* was classified as C1. The rest of the list ranged between B1 and B2. For non-emotional words, most of the words were, either B1 or B2. Only three words of the 16 words in the list belonged to the C2 level (*heritage*, *intellect* and *sequence*).

To link the classification of AWs to the recruited groups, it can be argued that the impact of word difficulty diverges, depending on the characteristics of each group. Starting with monolinguals, it is less likely to assume that they had difficulty when they encountered both emotional and non-emotional words. This is grounded because they are monolinguals of English, and the selected words are English words. It could be argued here that they would be able to understand the meaning of these words alongside scaling emotionality precisely. As a result, there were no implications for this factor among monolinguals. For L2 speakers, difficulty linked to these words was expected since they were not native speakers of English. The level of language proficiency for advanced L2 speakers in the current study ranged between 6.50 and above according to the IELTS score band. L2 speakers with this band score, according to CEFR, can understand words from A1 to C1. For beginners, the level of language proficiency ranged between 4 and 5 band scores in IELTS. L2 speakers with this language proficiency can identify words from A1 to B1.

Having stated the characteristics of both types of AWs in terms of word difficulty and the recruited samples, we can assume that there is a possibility that word difficulty was the reason behind non-systematicity in scaling non-emotional words. To identify its effect, let us consider the emerged patterns of emotionality in the current research, including both types of AWs. The three groups demonstrated several patterns. These patterns showed that emotional words were rated higher than non-emotional words, not only among monolinguals but also for L2 speakers. All three groups followed the same pattern in scaling emotional words even though there were some word variations in terms of difficulty according to CEFR. For monolinguals and advanced L2 speakers, this was expected, as their linguistic abilities allow them to not simply understand these words only but also to scale the level of emotionality precisely. The strength of rating these words did not reflect any difference between the three groups. On the other hand, the patterns of non-emotional words reflected unexpected patterns among beginners in rating these words higher than monolinguals. This occurred only among beginners, as both monolinguals and advanced L2 speakers did not differ in scaling the same set of words. Amongst L2 speakers, rating these words did not show any difference between beginners and advanced L2 speakers. Due to these complexities, such findings are open to several interpretations.

First, by looking at the frequency of both emotional and non-emotional words employed in rating the level of emotionality, we can see that the frequency of non-emotional words was higher than emotional words (Tables 5 & 6). As such, this might be attributed to the degree of frequency of these words. This factor has proven to be effective in different cognitive processes, particularly in the behaviours of words (Laufer, 1997; Quinlan & Dyson, 2008). Even though some words are classified as difficult based on their linguistic level, it is possible to assume that beginners came across these words many times, and that they could link these words to more emotional associations whenever they encounter them. Such an interpretation is also supported by how L2 speakers were recruited to collect data in the current research. Both L2 groups were recruited from the UK, as English is the language of communication. However, this could be generalised to all beginners.

The second interpretation of the emergence of this finding among beginners is related to the level of imageability linked to non-emotional words. We have indications from the literature that non-emotional words are affected by the level of concreteness and imageability (Altarriba *et al.,*1999). As touched on in Chapter Four, in the current study, the criteria for selecting non-emotional words included controlling the degree of concreteness. Such a selection was based on the characteristics of these words observed in the existing literature, as it was essential for methodological reasons to eliminate any impact of concreteness in the emerged patterns. Selecting non-emotional words with no limits in terms of the level of concreteness could affect rating the emotional intensity in an unsystematic way. A decision was made to control the selection of non-emotional words by controlling the level of concreteness. These words were expected to have a consistent pattern in rating emotionality across groups. However, this expectation was not observed in the outcome of the scale. Lack of consistency in rating emotionality among beginners leads us to question whether emotionality plays a role representing these words or not. In Chapter One, reference was made to types of images linked to AWs. It has been suggested that non-emotional words can be linked to images based on the literature (Altarriba *et al.,*1999; Paivio, 2006). The characteristics of these images vary, as images linked to non-emotional words are mainly symbolic. For example, a word, like *justice,* is linked to the image of a *frocked judge*. In the current research, the level of imageability was not controlled between fixed limitations, as it is difficult to examine the effect of all factors in one experiment. Therefore, this factor had a potential effect on the pattern of non-emotional words. This claim is based on two pieces of evidence here. First, evidence from the existing literature highlights the role of imageability in the representation of non-emotional words, even though these words do not have a physical existence in real life. Non-emotional words can be linked to images. One can argue that the impact of imageability becomes clear between beginners and monolinguals, as these two groups showed differences in scaling emotionality. It is assumed that the more L2 speakers acquire the language, the more they can identify and link non-emotional words to emotionality only. This is confirmed by finding no difference in rating these words between advanced and monolingual speakers in the current data. It is believed that beginners rated non-emotional based on two types of associations: emotional and symbolic images. This could justify why these words were rated high among this particular group only. By looking at Table 8 in Chapter Four, we can see why the level of imageability was not observed between beginners and advanced L2 speakers. It is assumed that rating non-emotional words was at the lowest level among monolinguals, then it increased gradually. Advanced speakers rated these words between monolinguals and beginners, as there was no difference between advanced L2 speakers and beginners. Between monolinguals and beginners, the difference was significantly clear. This is the second piece evidence on the role of images in rating non-emotional words. Such contrast in the behaviour of these words leads us to question if emotionality affects the representation of these words, or if there are other factors involved in the representation of these words. Taken together, the role of emotionality and imagery are available in these words. Findings from the current study, therefore, indicate that imageability, unexpectedly, plays a role in the representation of non-emotional words among L2 speakers.

7.3 Memory Patterns of Emotional and Non-emotional Words

In this section, there will be a discussion on the patterns that have emerged in STM and LTM tasks. Both characterise the behaviours of emotional and non-emotional words in two different memory stages. This section addresses the main findings of STM and LTM experiments, explaining the reasons behind differences/similarities in remembering emotional and non-emotional words. Specific details will be provided on the link between the memory patterns of STM, and whether they were similar/different to LTM. Further comparison will be made between various types of AWs and with other words, like CWs, with a physical existence[11]. Although these words are considered the control group in this thesis, the comparison helps in gaining a better understanding of the mechanism of storing them in STM and LTM compared to emotional and non-emotional words.

7.3.1 The way the recallability of abstract words was measured does matter

As discussed in Chapter Three, while there has been a range of literature which examined the recallability of both emotional and non-emotional words (Anooshian & Hertel, 1994; Ayçiçeği & Harris, 2004; Demaree *et al.*, 2004; Ferré, 2003; Ferré *et al.*, 2010, 2018), relatively little has examined the

[11] As stated in Chapter One, modelling the recallability of CWs in this study is based on linking the outcomes of STM and LTM tasks to Paivio and colleagues' dual representation of these words.

recallability of these words in STM and LTM among both monolinguals and L2 speakers. Since the current research aimed to provide insights into the memory patterns of these words on different samples. Of course, there were several patterns that showed some consensus in the lack of difference in remembering positive and negative emotional words and in remembering emotional words better than non-emotional words in LTM (Anooshian & Hertel, 1994; Ayçiçeği & Harris, 2004; Ferré *et al*., 2010, 2018). In terms of differences in the current study, the findings of the current research demonstrated differences with those from previous studies, including the recallability of emotional and non-emotional words in the STM task. In fact, the findings displayed some differences in showing no difference in remembering emotional and non-emotional words in STM task, for example. It was anticipated that emotional words, for example, would have a consistent memory pattern in terms of remembering these words better than non-emotional words, not simply in STM and LTM tasks but across groups. In the current study, there was a change in the task design in terms of the number of words to be remembered, and the time allotted to be remembered AWs. From a psychological perspective, particular principles are required to measure the recallability of words in any STM task. At first, the number of words required in each block should not exceed the capacity of storing items in the STM. Moreover, other restrictions, in the duration of words on the screen and the time allotted to remember emotional and non-emotional words, contributed to the emerging patterns in the STM task. Undoubtedly, it is the speaker's job to be highly attentive to see all the stimuli on the screen to keep as many words as possible in the STM to retrieve them later. When this control happened in the current task, it was possible to see that emotional and non-emotional words were remembered the same. As a result, the outcomes of the STM experiment did not observe any difference between emotional and non-emotional words. Therefore, it can be argued that the recallability of both types of AWs were affected by the characteristics of STM task in the emergence of these patterns. This unexpected pattern came as a surprise since this memory pattern occurred not just for emotional words but also for non-emotional words among L2 speakers. According to Cortis Mack *et al.* (2017), these issues relate to the task effect. To explain, this interpretation is based on the comparison between the current design compared to others in the literature. When speakers were given more time to remember, differences in remembrance between these words were observed. This finding was found in the literature (Anooshian & Hertel, 1994; Ayçiçeği & Harris, 2004; Ferré *et al.*, 2010, 2018). The behaviour of emotional words in these studies can support the task effect stated by Cortis Mack *et al.* (2017). Memory tasks designed in these studies were not rationalised based on the characteristics of STM.

At the word level, the semantic relatedness of word lists in the STM task is another factor which affected the lack of difference between remembering emotional and non-emotional words. According to Majerus and Van der Linden (2003), semantically related words show a high level of remembrance compared to lists with unrelated semantic characteristics. These two types of AWs have similar semantic features in referring to something intangible, but differ in relation to the type of emotions they describe. The list of emotional words in the current study combined negative and positive emotional words. The same applies to the list of non-emotional words, as these words combine have different levels of images. The emergence of this memory pattern could be affected by this factor too.

Another group of differences was observed in the outcomes of the LTM task. The results demonstrated different findings at the word level and across groups. These patterns can be classified as expected and unexpected. Starting with the expected patterns, the recallability of negative and positive emotional words did not differ among both groups. Also, CWs were remembered better than non-emotional words among both groups. On the other hand, there were some unexpected patterns amongst both groups. Among monolinguals, for instance, there was a difference between both types of emotional words compared to non-emotional words, and between emotional words and CWs. For beginners, there was no difference between remembering both types of emotional words compared to non-emotional words. Even between CWs and both types of emotional words, there was no significant difference in remembering.

Interpreting these findings requires investigating the impact of several factors on the recallability of emotional and non-emotional words in the LTM task. The lack of difference between remembering positive and negative emotional words emerged early on in the scale rating task for examining the level of emotionality linked to both types of emotional words. It was expected that this was based on controlling the degree of valence and arousal, reflected in the recallability of both types of emotional words in the LTM task. This interpretation is supported by the similarity in the classification positive and negative emotional words in terms of the level of difficulty. According to CEFR, positive and negative emotional words were classified as B except for two words only. Therefore, word difficulty has less effect on remembering emotional words, showing no difference in remembrance. However, when there is a change in the classification of words in terms of difficulty, we should expect an impact on memory patterns. The superiority in remembering CWs compared with non-emotional words among both groups is a good example. By looking at the CEFR classification for non-emotional and CWs, most non-emotional words were classified as B1 with four words classified as B2. CWs, according to CEFR, are classified as A except for the word *Jewel* in the LTM task. Based on this change, the results demonstrated a tendency to remember CWs better than non-emotional words among both groups. This was predictable, given the ease of remembering CWs compared with non-emotional words. At this point, it can be argued that the level of word difficulty explains the emergence of these memory patterns, as it is important to question whether the level of word difficulty is the reason for variations in remembering words in the LTM task.

To address this question, there is a need to link the level of word difficulty to all memory patterns in the LTM task. Looking at the classification of emotional and non-emotional words, most of these words were classified in the B level. By considering the unexpected results touched on earlier, several patterns could not be explained based on the level of word difficulty, such as the lack of difference between emotional words and CWs among monolinguals, and between both types of emotional words compared with non-emotional words. Even when comparing CWs and both types of emotional words, there was no significant difference in remembrance. It is, therefore, likely that the level of difficulty as a factor could have an impact on the recallability of words. However, sometimes the characteristics of the task may eliminate the effect of word difficulty. In general, remembering words classified as A, for example, should be easier when compared with words in B or C level. At this point, there seems to be a task effect again in the way emotional and non-emotional words are patterned in LTM task. The effect is related to the availability of two elements: the learning phase and contextual associations. All of these elements helped in increasing the number of remembered words. In fact, the learning phase facilitated an explanation of the meaning of these words if the speakers had not come across these words before. This eliminated the effect of word difficulty between both types of AWs and CWs.

The availability of contextual associations in the LTM task helped monolinguals, particularly in remembering the exact word since their memory patterns showed no differences in remembering, either emotional or non-emotional words. This interpretation is supported by the behaviours of emotional and non-emotional when compared with CWs. Even the recallability of CWs, which are different from AWs, gives us a reason to assume that the characteristics of the task encouraged these memory patterns to appear. Having realised the importance of contextual association in measuring the recallability of emotional and non-emotional words, one can argue here that this factor is responsible for the dominance in remembering positive emotional words compared to non-emotional words among L2 speakers in the LTM task. The difference emerged because of the intensity linked to positive emotional words. The role of context in retaining words in the LTM, according to Tulving (1972), was effective, as words do not stand alone in the memory. They are linked by forming networks and associations (Kiss, 1968; Paivio, 1971a; Wilks & Meara, 2007). The availability of different associations makes the recallability of these words easier. The characteristics of these associations vary and can involve both verbal and visual associations.

Such findings assist in addressing the first mini research question relating to experiments 2 and 4 on the recallability of emotional and non-emotional words. This question, which examined how do L2 speakers remember emotional and non-emotional words in STM, can be answered now by looking at how L2 speakers remembered emotional and non-emotional words in the STM task. The outcomes indicated similarities among monolinguals and L2 speakers in relation to remembering the two types of AWs. This leads us to consider the first hypothesis of experiment (2) in Chapter Five, where emotional words are hypothesised to be remembered better than non-emotional words in the STM task. The findings did not show any difference between these words regarding recallability in the STM task, and this hypothesis could not be proven. The null hypothesis, instead, could be accepted based on the outcome of the testing environment. As for the behaviours of emotional and non-emotional words in LTM task, the difference between these words is attributed to various factors, like the semantic characteristics of these words, the level of word difficulty and the design of the LTM task. Before running the LTM experiment, it was believed that these two types of AWs would be remembered differently, and this research hypothesis could be partially accepted, as positive emotional words were remembered better than non-emotional words among speakers with a modest level of linguistic abilities only. Therefore, the first mini research question (1) in Chapter Six, which focused on how L2 speakers remember emotional and non-emotional words in LTM, could be answered based on the dominance of remembering positive emotional words better than non-emotional words.

7.3.2 Primacy effect in remembering abstract words in the short-term memory

The current research observed the primacy effect in examining the STM task. This finding addresses the third research question (3) in Chapter Five which questioned the existence of primacy or recency effect in remembering emotional and non-emotional words in the STM task. The current research addressed this issue, and provided insights into this important area. By looking at the previous work, there was an emphasis on the order of remembering words in STM. This emphasis can be classified into two parts. The first part examined the effect of primacy/recency in the order of remembering items in general. While there has been a variety of literature which has explored whether items in the initial string are remembered better compared to items in the final position (Baddeley *et al.*, 2009). There has been a specific debate on the effect of primacy/recency in remembering words, representing the second part of the literature. This part demonstrated the effect of primacy on the recallability of various types of AWs in particular (Demaree *et al.*, 2004; Yoo &

Kaushanskaya, 2016). This is important because this feature in identifying items - including words - follows a particular order by better remembering AWs in the initial position or in the final position. Each of them reflects a different understanding of how these items behave in the memory.

The findings from the current research showed some similarities with the existing literature, particularly in highlighting the primacy effect in remembering emotional and non-emotional words among monolingual speakers (Demaree *et al.*, 2004; Yoo & Kaushanskaya, 2016). This pattern of remembering words in the initial string was also observed in relation to remembering CWs among both monolinguals and L2 speakers. In contrast, remembering emotional and non-emotional words among L2 speakers suggested that this pattern could not be applied to L2 speakers. There was no effect in remembering these words in the initial position. The similarity between the existing literature and the current study in observing the primacy effect in emotional and non-emotional needs to be highlighted, as the current research drew it from monolinguals. This is suggested in what way memory works in remembering these words. In terms of the markedness of the primacy effect, what can be learned here is that monolinguals keep rehearsing words shown at the beginning of the string. This is to avoid forgetting them when the string is over. Subsequently, words at the beginning of the string were retained in the speaker's STM. As a result of the primacy effect, words that occurred in the final position were impacted. The lack of recency effect in remembering emotional and non-emotional words in the final string indicates that these words vanish quickly from memory. Speakers tried to keep their memory active by rehearsing the words demonstrated earlier. By doing that, words in the final string were forgotten. Put differently, when there is a primacy effect in remembering words, the recency effect will be affected. This pattern in remembering words in the initial position was observed among monolinguals in remembering CWs too. These words were continuously remembered in the initial position better than the final position. This tells us that even with changes to word type in the STM task, the primacy effect was still observed among monolinguals.

There were also important distinctions between existing research and the findings from the current study. There was a lack of primacy effect in remembering emotional and non-emotional words among L2 speakers in the current study. It was expected before running the experiment that there would be a primacy effect among L2 speakers. This speculation increased when the primacy effect was observed among monolinguals, and no recency effect was observed. Such contrasting findings imply that the order of recallability of these words among L2 speakers follows a different pattern, showing no effect for primacy or recency. Observing no difference in remembering emotional and non-emotional words among L2 speakers in the initial string suggests that L2 speakers had a consistent pattern in STM for both types of AWs among L2 speakers. What can be understood here in reducing the primacy effect is caused by the attention speakers give to remembering these words. Participants were required to remember as many words as they could in a time limit of 30 seconds. Due to the semantic characteristics of emotional and non-emotional words in referring to something that is not tangible, speakers had to concentrate on finding any association which helps them to retain emotional or non-emotional words in their memory. Finding such associations for AWs was more difficult than for CWs, especially when presented in another language. Moreover, the time of word presentation and remember complicated the matter. Speakers had a limited time to retrieve these words.

It is less likely to be assumed, based on the outcome of STM task, that speakers link both types of AWs to any emotional associations, especially words in the initial string. This interpretation is supported by how these words work among monolinguals. The primacy effect was observed, as these words were presented in English to monolingual speakers of English. Monolinguals seem to find it easier to link the presented emotional words to emotional associations in a timely constrained condition compared to L2 speakers. Based on this, the primacy effect was observed among monolinguals not L2 speakers. As for the lack of difference in remembering both types of AWs in the final position, this effect was not expected since the list of words is shorter than 9.

Providing items less than this number, it is more likely the effect of recency would not be marked. In the current data, this effect was not observed. There was a lack of recency effect in the recallability of emotional and non-emotional words among monolinguals. Not only this finding was observed in the recallability of emotional and non-emotional words among L2 speakers but also in CWs. This indicates that the recallability of emotional and non-emotional words reflects a consistent pattern in STM in the order of remembering these words. This tells us that the recallability of L2 speakers for emotional and non-emotional words has followed one pattern, and it is shared with speakers who can only speak one language.

The way that both emotional and non-emotional words are remembered represented the essence to answer the third question in the experiment (3) in Chapter Five. The primacy effect was found in the outcome of STM task followed Baddeley's model (2009). This effect was realised in the recallability of emotional and non-emotional words among monolinguals alongside the recallability of CWs among both types of the lexicon. Therefore, the third research hypothesis (3) in the STM task on hypothesising the impact of primacy in the recallability of AWs could not be proved among L2 speakers only. The outcome of STM task among monolinguals could prove this hypothesis.

7.3.3 Imagery effect in the short-term and long-term memory

After studying the memory patterns of emotional and non-emotional words in two different memory types, the recallability of CWs was studied in both STM and LTM tasks. Although AWs and CWs are different semantically in their meaning, and cognitively in the way they are represented, both monolinguals and L2 speakers were able to remember CWs in STM better than emotional and non-emotional words. This tells us that CWs continue to have a steady pattern in these memory types. This shows some consensus with the literature CWs are better remembered (Paivio, 1971a, 1973, 1991). Such confirmation is important here to be emphasised since it showed agreement with the existing literature in terms of the dominance of remembering CWs in STM which is associated with the role of imagery (Altarriba *et al.*, 1999). In the current study, examining these words was important even though they were considered a control group, as comparing the recallability of these words - which have a high level of associated imagery - with words with a high level of emotionality, is crucial for determining the strength of imagery in memory retention compared to emotionality.

This consensus between the literature and the current finding in the recallability of CWs can be interpreted based on the characteristics of these words according to both CEFR and the dual representation of CWs in DCM. CWs included in the STM and LTM tasks were classified in the A category, except the word *Eagle* in the STM task and *Jewel* in the LTM task. Judging according to word difficulty, we can see that CWs were easier to remember. These words belonging to the A category are easier when compared to other categories. The results of the STM task between monolinguals and L2 groups confirmed this claim. In the LTM task, the same pattern was observed in relation to the superiority of remembering CWs compared with non-emotional words only. What we can learn about word difficulty here is that when words vary in their CEFR classification, the impact of this factor occurs. Variations between CWs and emotional and non-emotional words support this conclusion. At the group level, superiority in remembering CWs was observed among all groups. Regardless of the way CWs have been examined or the characteristics of the recruited informants, these words were constantly better remembered in the STM and LTM tasks.

Existing literature has touched on the dual representation of CWs, providing dual representation of these words. As stated earlier in Chapter One, Paivio and colleagues' DCM theorised on the dual representation of CWs, which led CWs being better remembered than AWs. Since the design of the STM task in the current study followed Paivio and Csapo (1969), the findings asserted the role of imagery in remembering CWs better than both types of AWs. Although the pattern that emerged in Paivio and Csapo (1969) was elicited from monolinguals and bilinguals, an important contribution of the current research is in observing the same pattern in relation to CWs but among L2 speakers with various linguistic abilities. For CWs, associated imagery is shown to help in remembering these words better and facilitates the recallability of these words. Imagery represents one level of representation in addition to the verbal level which belongs to all words. The characteristics of a word, like *school,* reflect a mental image(s) whenever this word is uttered or seen. This makes the recallability of these words easy, suggesting that there is one shared imagery system since there was no difference in remembering these words across groups. This imagery system is called referent images, shared by the two languages. Even if the verbal representation (language) differs, the recallability of CWs was still marked. This system facilitates the recallability of these words in STM and LTM tasks. Moreover, the superiority in remembering CWs can be attributed to the semantic characteristics of these words since they refer to something which human senses can observe. These findings came to answer the second mini research question (2) raised in Chapters Five and Six on how the recallability of CWs differs among L2 and monolingual speakers. CWs showed a consistent pattern in remembering these words among different groups in different memory types, varying in nature in their designs and characteristics. Thus, the findings from the STM and LTM tasks confirm the two hypotheses relating to the superiority of remembering CWs in Chapters Five and Six.

7.4 Hypotheses Revisited on the Relation between Emotional Intensity and the Recallability of Abstract Words

After discussing the findings of emotionality rate and the recallability of various types of AWs among two different types of the lexicon, it is needed to consider revisiting the hypotheses. In Section 1.2.5, there was an explanation on the hypotheses of the thesis. It was assumed that the distribution of emotionality varies at the word level between emotional and non-emotional words on the one hand, and between L2 speakers and monolinguals on the other. The emergence of these patterns in the thesis leads us to believe that the level of emotionality was different at the word level of AWs with no difference amongst L2 speakers. Both emotional and non-emotional words were rated with no significant difference among L2 speakers. Even the strength of rating these words was similar among L2 speakers. This indicates that the part of the first research hypothesis (1) stated in Chapter One needs revisiting. The level of language proficiency was not observed in scaling AWs among L2 speakers. The failure to observe the impact of language proficiency related to emotionality being a psychological factor. It impacts AWs through the linking of emotional associations with words; something acquired in early childhood, and it is formulated first in the speaker's L1. As a result, the impact of the speaker's L2 is limited and ineffective, especially when the level of language proficiency opens to changes. In this case, L2 speakers rely mainly on the emotionality established from their L1 and rate AWs accordingly. The similarity between monolinguals and L2 speakers at this point supports this explanation. The recallability of AWs in STM and LTM tasks showed a need to consider the second research hypothesis (2) in Section 1.2.5. Originally, it was assumed that emotional words are remembered better compared to non-emotional words in both STM and LTM among L2 speakers. The findings emerged from running memory tasks in the thesis indicated that the recallability of emotional and non-emotional words is affected by many factors. Assuming that the emotional intensity perceived with AWs was the cause behind remembering emotional words better is not well-justified. Each one of these memory types has characteristics which affect how emotional and non-emotional words are remembered. The characteristics of STM task did not allow observing the difference among L2 speakers. Speakers had to remember as many words as possible in a time-constrained manner, so their attention was focused on remembering all the words. Similar to the patterns of emotional intensity, the performance of L2 speakers reflected the same pattern in remembering emotional and non-emotional words similarly, with no variations between competent and modest speakers. Due to these complexities, the recallability of various types of AWs is affected by many factors, like the design of memory tasks, semantic characteristics, word difficulty and the level of emotional intensity.

7.5 Limitations in Studying Abstract Words

Whilst this research has provided important insights on the recallability of AWs, it has not been possible to comprehensively address all of the processes involved. Furthermore, additional limitations can be identified at both the theoretical and methodological levels which are important to clarify. Highlighting these limitations is crucial for further understanding AWs.

7.5.1 Theoretical limitations

Chapter One discussed the theoretical models used in the thesis, including those of Kousta and Vigliocco (2010) and Baddeley *et al.* (2009). The role of these models was central in consolidating the argument of the thesis and interpreting the emerged patterns either in scale rating or memory retention. Each has been used to interpret the behaviours of AWs under a specific testing condition. Kousta and Vigliocco’s model (2010), for instance, contributes a set of ideas about emotional words, and how the semantic features of these words provide links with more emotional associations. In terms of examining the order of remembering AWs in the STM task, Baddeley *et al.’s* work (2009) was employed, helping to explain the reasons behind the emergence of the primacy effect in remembering words. No other theoretical models have been used in interpreting the results of AWs in the current study. Taken together, these models proved to be the most suitable for the current research study, and enabled valuable insights to be drawn from the findings. This could be considered as one of the limitations of the current research.

7.5.2 Methodological limitations

At the methodological level, there was another set of limitations related to the approach of examining AWs and the selection of AWs. There was an explanation of the approach being used to study these words in Chapter One. By looking at the approach of measuring the recallability of these words in the current thesis, we can see that the research examined the effect of emotionality on the recallability of emotional and non-emotional words at the retention and retrieval stages. The patterns emerged for these words were restricted to these two stages of memory only. Another limitation can be highlighted in relation to the selection of AWs. The AWs chosen for the study were singular nouns on the assumption that different grammatical units of AWs (e.g., *adjectives* and *verbs*) may show different memory patterns. This selection was based on the generalisation found in the literature that mixing nouns and adjectives can elicit false generalisable results. These limitations aimed to control the testing environment of AWs according to well-defined criteria for selecting AWs. Choosing a list, which combines different grammatical units, may affect the findings of the study. All in all, this limitation as well as others helped in studying the organisation of English AWs with the effect of emotional intensity. These choices were still the most suitable for the study given what the study aimed to explore

7.6 Contributions of the Study

Previous research has focused on word type which has led to the omission of consideration of groups. As the current study has shown, this is an important oversight. The three groups have significant differences in scaling emotionality concerning AWs. Therefore, it has been essential to address this omission and to identify the contributions to knowledge that can be gained from adopting this approach. The contribution of the study is evident at different levels: the theoretical, methodological and tangible levels.

7.6.1 Theoretical contributions

Whilst the thesis focused on AWs and the way they are remembered, two significant theoretical contributions emerged from the outcomes of the study. The patterns identified, for instance, contribute to the models involved in interpreting the establishment of emotionality and the recallability of words. To be clear, the current study suggests:

- The extendibility of the emotionality model, which was included in the study, on L2 speakers with different linguistic abilities. This issue is considered one of the limitations of the model, as the model relied mostly on monolinguals and bilinguals. Between these two samples, there is a need to be clear in determining the existence of this model. Proving its existence in the behaviours of emotional words among L2 speakers whose linguistic abilities varied is considered a key contribution to the existing work.

In terms of memory patterns appeared in this study, CWs were included to widen the comparison between words with no physical existence, such as emotional and non-emotional words, on the one hand, and words with physical existence in life on the other. As shown in Chapter One, the outcomes indicate the dominance in remembering CWs compared to emotional and non-emotional words. This result was expected from the beginning, especially for monolinguals. However, what this study also provides is suggesting the same results among different samples. We have seen that one of the limitations of DCM was the existence of this dual representation among L2 speakers. As such, this study underlines:

- The extendibility of the DCM in the recallability of CWs among L2 speakers. There is a gap in the existing work of this model on the possibility of including L2 speakers in the way DCM conceptualises CWs in the memory. Most research on these words used different samples, yet evidence of the use of this model amongst L2 speakers was lacking.

7.6.2 Methodological contributions

Another level of contribution reflected in this study relates to methodology. The study followed sequential steps in creating the best environment to study AWs. The proposed context of studying AWs in the research was unique since it focused on testing the effect of emotionality on these words among L2 speakers whose abilities have been properly assessed. This approach of testing AWs improved our understanding of AWs, and the way they are organised in the lexicon. Such an approach may prove influential to future research on AWs.

7.6.3 Tangible contributions

This research provided key contributions in understanding of the patterns of AWs. These contributions were witnessed in the establishment of emotionality of AWs on one hand and the mechanism of storing these words with various levels of emotionality in the memory on the other. The study offered valuable insights into AWs in terms characterising the level of emotionality by:

- Linking the establishment of emotionality to the semantic features of AWs. This factor enables the acquisition of more emotional associations with emotional words regardless of the type of the lexicon. The meaning of non-emotional words led to rate differently in terms of emotional intensity, making these words similar to emotional words among L2 speakers; and

- Demonstrating that the speakers' linguistic ability does not influence emotionality as a psychological factor. The role of language proficiency in causing no effect in establishing emotionality came to contribute to this conclusion about AWs when it was studied among L2 speakers.

At the memory level, the way AWs were measured affects the emergence of memory patterns in STM and LTM tasks. The patterns of AWs changed when the type of memory changed. This contribution is observed in the following:

- The lack of difference between emotional and non-emotional words in STM, where there were time restrictions to demonstrate and remember words. Even with the high level of emotional intensity found in emotional words, this feature did not affect the recallability of these words. This pattern was seen in both types of the lexicon. Even among L2 speakers, variations did not affect the recallability of both types of AWs; and
- The primacy effect was observed in remembering emotional and non-emotional words among monolinguals in STM task.

Taken together, these contributions in the establishment of emotionality as well as the emerged patterns in memory tasks came to draw a wider view of AWs to determine, whether there is an impact of emotionality in the recallability of AWs. This study encourages more emphasis on AWs, with a focus on developing this area of research to address the knowledge gap in the literature. The current research indicated an association between memory performance and emotionality, but this association is affected by linguistic and psychological factors. From this, we can argue that emotional intensity enhances memory capacity in storing AWs in the memory, but the impact of this factor depends on the type of AWs as well as the nature of the memory task. This study contributed to studies on AWs which had not yet been examined from this perspective. These results are one of the most interesting contributions of the study which happened as a result of extensive research into these types of AWs and experiments conducted on the selected samples. Confidently, the outcomes of this study help to provide a better understanding of how the lexicon interacts with these words.

7.7 Implications of the Study

The outcomes of the study enhance our understanding of AWs, especially in the L2 context, as the existing literature on AWs has placed little emphasis on these words. The approach adopted in the study contributes to the fields of lexical semantics, psycholinguistics L2 acquisition and psychology, and the outcomes offer a broader and more precise perspective of AWs. It is important to declare that the patterns of AWs were inspiring. This is based on the fact that unexpected results emerged in the current study. The lexicon deals with these words differently, depending on many factors, like emotionality and bilingualism. Since this thesis mainly focused on the L2 lexicon, more patterns emerged with a particular reference to the level of emotionality. The results imply that the processing of these words in the lexicon is complicated regarding word storage. Both factors are involved in processing these words in the lexicon, especially among L2 speakers.

To be more specific, the results of this study have practical implications. First, examining these words in a series of testing conditions using different theoretical models revealed that AWs in lexical semantics should be given more attention. The constant comparisons between these words and CWs in other studies have only occasionally been well-reasoned. AWs cannot be directly compared to other words that can be related to physical objects. AWs have a more advanced level in terms of their meanings, references and features. It is hard to assume that emotional and non-emotional words, as different types of AWs, have the same semantic features. Although a few studies have acknowledged this issue, the vast majority of the literature still needs to consider the semantic features of these words. The results of this study have implications not only from a lexical semantics standpoint but also from a psycholinguistics perspective. The behaviours of these words have been presented in this study, highlighting the role of emotionality in the organisation of AWs in the lexicon. By examining the role of emotionality, we can state that many factors affect the representation of AWs. For psycholinguists, the outcomes of these experiments help to re-evaluate the role of emotional intensity in processing language overall and AWs in particular. While the current study examined only emotional and non-emotional words, it is worth considering emotionality in different linguistic inputs. The role of psychological factors - like emotionality - on the linguistic structure of the lexicon among monolingual or bilingual speakers is important to consider.

The implications of the study can be extended to L2 acquisition studies. Applying the model of embodied emotionality to AWs in the L2 context helps to explain how these words are processed across the two levels of language proficiency. Therefore, in this study, AWs showed different behaviours among L2 speakers, and they are stored differently in the bilingual memory. This conclusion may change how L2 specialists think about different memory patterns when storing these words. The recallability of various types of AWs differs and, therefore, cannot be generalised to all memory conditions. At the same time, the level of language proficiency matters in some memory conditions. More patterns are expected to emerge with different memory conditions at different L2 levels.

In psychology, psychologists need to consider the semantic characteristics of words before making any generalisations. Even the confounding variables need to be considered. Patterns of abstract nouns will be different from adjectives and verbs. Chapters One and Three stated that relying on ready-made word lists in the literature may lead to inaccurate results with false conclusions and interpretations. It not easy, therefore, to arrive at any conclusion about the memory patterns of AWs without considering the semantic features of these words before running any psychological studies.

7.7.1 Implications in teaching various abstract words to L2 speakers

The findings of this study have important implications for teaching, particularly for language pedagogy, and the way these words should be introduced to L2 learners. After demonstrating the differences between various types of AWs semantically and cognitively in reflecting different levels of emotionality, it can be argued that L2 language instructors should introduce AWs with more caution. Non-emotional words should not be presented with emotional words, as non-emotional words are most likely to become passive vocabulary or be forgotten easily among L2 learners. These words should be introduced with context since this plays a crucial role in enabling them to be stored in LTM. L2 learners should also be asked to use these words in different contexts. By doing so, more associations can be established with AWs, helping them to be stored when needed in the LTM. This strategy for both types of AWs, but it may be more effective with non-emotional words. To store various types of AWs in LTM, L2 language instructors can diversify the type of associations with AWs, not only with simple static pictures but with poetic lines, famous lyrics, symbolic signs and moving pictures. If a particular AW has been introduced with different associations, either verbally or visually, it is likely to be stored permanently in the memory. While introducing a particular word with different types of associations may be challenging and time-consuming to L2 instructors, the characteristics of AWs force us to deal with them differently. From the participants' perspective, there are diverse ways of introducing these words to suit different L2 speakers with individual differences. Some learners prefer to link newly learned words with static pictures. Others prefer to link words to lyrics and songs, as these contextual associations are already stored in the learner's memory. If learners have no previous associations or cannot establish new associations with AWs, they can be advised to link newly learned words to old memories, experiential knowledge or affective states in their native language. For emotional words, this will be straightforward since the semantic meaning of these words helps. However, for non-emotional words, L2 learners may struggle to link these words to different levels of emotional intensity. Language instructors should bear in mind this challenge faced by L2 speakers. These insights for teaching AWs suggest that language instructors might benefit from starting with cognate AWs between the learners' L1 and L2 (e.g., *noble in English* – نبيل in Arabic). Learners tend to retrieve AWs from memory with the same associations in their L1. Overall, the intensity of presenting AWs in different ways, linking them to old memories or experiential knowledge could be effective in changing AWs from being passive to active. This change is observed in asking L2 learners to use both types of AWs frequently in ongoing communication. These implications do not have to be a part of a textbook. According to Milton (2009), good language instructors can expand teaching environments, and provide a rich lexical environment to examine the usability of words by providing tasks to help learners store and use AWs.

7.8 Summary

This chapter has discussed various distributions of emotionality within AWs. It analysed the potential impact of emotionality on the recallability of emotional and non-emotional words. The outcomes of the four experiments solidified the argument stated earlier at the beginning of the thesis about the complexity of AWs in terms of storing these words in the memory differently. The emergence of these patterns in rating the emotional intensity, and the recallability of emotional and non-emotional words are considered to be a key contribution to human knowledge at different levels. Emotionality as a psychological factor, for instance, does not solely influence recallability. There are other factors that affect how AWs are stored in STM and LTM.

Chapter 8

Conclusions and Future Recommendations

8.1 Introduction

This chapter provides a summary of the key findings from the study for both monolingual and L2 groups. The chapter hints at some of the areas that future research into AWs can explore. The chapter concludes by returning to the core research questions for the study, and how they have been addressed.

8.2 Summary of the Study

In Chapter One, there was an outline of the essential parts of studying AWs in this thesis. These parts included the background of the thesis, the statement and the significance of studying these words. Context of the study, as well as, the research questions and hypotheses were provided to explain how the impact of emotionality affected the recallability of AWs in L2 lexicon. The conceptual framework of the study was included too in the First Chapter. In Chapters Two and Three, there was a detailed discussion on the identification of emotionality as a psychological factor and memory as a cognitive process respectively in relation to AWs. Starting with examining emotionality, Chapter Two examined emotionality as a concept, and how it relates to the language at the sentence and word levels. The characteristics of emotional intensity were explained in terms of emotional associations linked to AWs. Further details were provided on the distribution of emotionality of AWs among monolinguals and L2 speakers, and whether the level of emotionality varies. The process of storing words in the memory represents the essence of Chapter Three, including features and restrictions between memory types. The discussion in the chapter was centralised primarily around the characteristics of STM and LTM, as there was a link to these two types of memory to examine the recallability of AWs. Both Chapter Two and Three ended with examining how various types of AWs were studied in terms of the distribution of emotionality and storing these words in the memory at the retention stage among both monolinguals and L2 speakers.

The behaviours of AWs in the current study in three testing conditions in Chapters Four, Five and Six led us to believe that the recallability of these words is different among monolinguals and L2 speakers. The recallability of these words is affected by many factors. While some of these factors, such as semantic characteristics, word difficulty, level of emotionality and level of imagery have proved to be effective, others - like the level of language proficiency among L2 speakers - have been found less effective. One of the most important factors is the semantic characteristics of AWs, as this is prominent in how memory retains and retrieves emotional and non-emotional words. However, the impact of this factor does not affect the recallability of AWs in a timely controlled condition, like the STM task. Based on this classification, both emotional and non-emotional words had different memory patterns as shown in Chapters Five and Six. They can be summarised as:

- Emotional and non-emotional words behave similarly with no difference among L2 speakers in the STM task;
- The primacy effect in the STM task was observed in remembering emotional and non-emotional words among monolinguals;
- The recency effect was not observed in remembering emotional and non-emotional words among both monolinguals and L2 speakers;
- The level of language proficiency did not affect how various types of AWs were remembered among L2 speakers in STM task;
- Differences in remembering emotional words were only shown in the superiority of memory in the LTM task not in STM;

- There were some similarities and differences in remembering these words among L1 and L2 speakers;
- At the word level, there was a significant difference in remembering positive emotional words and non-emotional words in the LTM task, unlike negative emotional words;
- At the group level, there was a significant difference among beginners in remembering positive and non-emotional words in the LTM task; and
- No significant difference was observed between positive and negative emotional words in the LTM task among both monolinguals and beginners for a methodological reason caused by controlling the level of arousal and valence.

As this study measured the level of emotional intensity in Chapter Four, different patterns emerged in that process which can be listed as follows:

- AWs differ at the level of emotionality, varying between emotional and non-emotional words. Levelling emotional intensity was based mainly on the semantic characteristics of AWs, as emotional words have a higher level of emotionality than non-emotional words. Differences, however, emerged between these words at the word level rather than at the level of language proficiency;

- Similar to the outcome of LTM task, variations within emotional words could not be observed for a methodological reason caused by controlling the level of arousal and valence;
- Non-emotional words showed unexpected patterns among L2 beginners in scaling these words higher than emotional words; and
- In terms of the strength of emotionality, no difference was observed among monolinguals and L2 speakers in rating these words.

The analyses of the three tasks across the two levels of L2 speakers concluded that L2 speakers vary in rating AWs in terms of emotional intensity and in storing them in STM and LTM. The findings of the current study have contributed to knowledge by addressing this gap in studying AWs.

8.3 Recommendations and Potential Extensions of the Study

Whilst this research has contributed key insights into AWs, there are many areas still need to be explored. In psycholinguistic studies, further studies could compare the representation of abstract nouns to adjectives in the lexicon to different interactions of emotionality within AWs. It would be interesting to look at the level of emotionality between abstract verbs and adjectives and, link it to the recallability of these words. More work is also encouraged on AWs with different numbers of syllables. As a reflection of previous work in the literature, linking the representation of AWs to memory provides a broader horizon for AWs. One highly recommended area to study within AWs is the variation of abstract emotional and emotionally-laden words (e.g., *cancer*). Abstract emotional words have different semantic characteristics. They are subcategorised as there are emotional descriptors (e.g., *happiness*) and emotional arousers (e.g., *death*). Other abstract emotional words are still defined as emotional words, but are neither descriptors nor arousers. Therefore, the lexicon processes these words and stores them in the memory differently. It would be helpful to conduct a comparative study between monolinguals and L2 speakers on the variation of emotional words alone. Furthermore, it is recommended that either emotional or non-emotional words within sentences or phrases based on the word's emotionality and the availability of context are further explored. Some examples in the literature focused on neutral expressions, terms of endearment, insults, reprimands and taboo words within phrases (e.g., Caldwell-Harris *et al*., 2011). Words that are tested should be provided with a context and rated for their level of emotional intensity. Overall, investigations into the representation of emotional words in the L2 lexicon using context are still lacking. The level of emotionality of AWs in one task should be further measured in two languages simultaneously to draw a more comprehensive comparison.

At the task level, it would be encouraging to increase the number of every type of AW in the STM task. The investigation could be made to study variation in emotional words by testing the recallability of positive and negative emotions separately without randomisation between word types. The suggested classification of words in STM can be positive, negative or non-emotional, and each group consists of the same number of words. By showing no difference between emotional and non-emotional words in the results for STM task among L2 speakers, it is recommended that these two types of AWs are examined separately and the results compared. This selection can also examine the effect of emotionality on the semantic categorisation of AWs. The homogeneity/heterogeneity of word blocks in memory tasks, for example, may affect the recallability of words (Poirier & Saint-Aubin, 1995). The same applies to LTM task, making the number of all types of AWs even. Furthermore, the participants could be classified based on their level of intelligence, as the speakers' intelligence can be assessed by the *intelligence quotient (*IQ). This factor also affects the outcomes for memory patterns, especially in measuring STM (Cowan, 2008). Regarding the characteristics of the participants, more work exploring the role of emotionality by gender is necessary which was not included in the current work. From a neurolinguistic standpoint, the structure of the lexicon differs between males and females. Therefore, future work could investigate the organisation of these words across genders. Moreover, the study could be replicated with left-handed participants to determine whether the change in this variable affects the organisation of AWs. By examining L1 and L2 together, more patterns are expected to emerge concerning the emotionality of AWs and in memory patterns. It is desirable to study the representation of AWs among people of different ages, and make a broader comparison between groups. Furthermore, studying these words amongst multilingual speakers may provide additional insights into the patterns of these words.

At the theoretical level, further work is needed to study the structure of the L2 lexicon, concerning AWs based on a different model. Despite the important results that emerged in this study, studying AWs in the L2 context by testing the organisation of AWs using different theoretical models may facilitate understanding further. The representation of AWs can also be evaluated from the perspectives of psychology and neurolinguistics. Future work in psychology on these words could examine different types of memory, measuring the stages involved in storing AWs. Further work could be done to measure the retrieval process of these words from both lexicon types using RT software. In neurolinguistics, examining the brain stimulation of emotional and non-emotional words with specific reference to the areas involved in storing these words, using MRI or electroencephalography is promising. The elicited data of the current study is inspiring to keep working on AWs. Certainly, this field will offer a different understanding of AWs from a different standpoint. Using varied methodologies and research designs, it will be another motivational journey to learn about these words in a different testing environment.

8.4 Conclusion

This thesis aimed to identify the patterns of emotionally charged words in the memory, and to find out whether this feature impacts the recallability of these words among two different types of lexicon. The distribution of emotionality represents the first research question of the thesis (1) which examined how L2 speakers of English rate various types of AWs in terms of their emotionality compared to monolinguals. The patterns that emerged proved the first part of the research hypothesis (1) which hypothesised that there are variations in scaling the emotionality of AWs at the word level. Emotional words are rated higher than non-emotional words. Both L1 and L2 speakers showed no difference in rating emotionality between emotional and non-emotional words with no effect on language proficiency (see Section 7.4 for revisiting hypotheses). L2 speakers rely mainly on the emotionality established from their L1, and rate various type of AWs accordingly. Even the strength of rating these words was similar among L2 groups. The similarity between monolinguals and L2 speakers at this point supports this claim. Clearly, emotionality is a psychological factor which affects AWs by linking emotional associations to words. This factor is acquired from early childhood, and it is formulated first in the speaker's L1. As a result, the impact of the speaker's L2 is limited and ineffective, especially when the level of language proficiency opens to changes. These findings can answer the first research question of the thesis (1) which questioned how do L2 speakers rate emotional and non-emotional words among monolinguals and L2 speakers.

Following the same conclusion in rating emotional words higher than non-emotional, it was expected that there would be a superiority in remembering emotional words better than non-emotional words. The patterns emerged from running the STM task showed no difference in remembering emotional and non-emotional words at the group level. However, the current study demonstrated different patterns of remembrance in terms of memory retention. These patterns answer the second research question of the thesis (2) which questioned the underlying memory patterns of emotional and non-emotional words at the retention stage among monolinguals and L2 speakers. The findings of STM and LTM tasks showed different memory patterns. Some of these memory patterns are shared between emotional and non-emotional words on one hand, and between monolinguals and L2 speakers on the other. At the same time, there are some differences in remembering emotional and non-emotional words. The emergence of these patterns is attributed to many factors. This leads us to question if the level of emotional intensity is responsible for the emergence of these patterns only. It was essential to address this supposition, as it represented the third research question (3) stated in Chapter One. The difference, however, in rating emotional words higher than non-emotional words is attributed to the semantic characteristics of emotional words. These characteristics allow associations, like delightful memories, past experiences and affective statues to be linked to these words in the lexicon. On the other hand, non-emotional words have less or no emotional associations, as their semantic characteristics do not allow them to link emotional associations. Therefore, the distribution of emotionality in the current research supports this interpretation.

How memory types were designed in the current research may affect the markedness of emotionality in remembering emotional better. We have demonstrated various examples in the thesis on the effect of the design on the emergence of memory patterns. This indicates that memory is a complex cognitive process affected by many factors, and one of these factors is the level of emotionality. Other factors, stated earlier in this chapter affect the recallability of both types of AWs, like the semantic features of AWs alongside the level of word difficulty. Therefore, some unexpected patterns remain in the current study, like the lack of difference between emotional and non-emotional words in the STM task, even though these words varied in terms of emotional intensity. Based on this, it is less sensible to assume that emotionality affects the recallability of emotional words at any stage. The current study showed the memory patterns in STM and LTM tasks, and explained the impact of memory designs and the level of word difficulty on the emergence of memory patterns. Even the level of language proficiency did not affect the recallability of AWs in the STM in particular. Overall, these various patterns assist in addressing the third research question (3) by arguing that the impact of emotionality could not be observed in the outcomes of all memory patterns at the retention stage. It is hard to accept the second research hypothesis (2) in proving the role of emotionality in the recallability of various types of AWs. Emotionality as a non-linguistic factor may affect the recallability of words like AWs, but other factors could affect the recallability of these words regardless of the type of AWs or the type of the lexicon. Taken together, the establishment of emotionality as a psychological process and memory retrieval as a cognitive procedure are two independent processes which are unaffected by the speakers' linguistic abilities.

List of Abbreviations and Symbols

AWs	Abstract words
α	Cronbach's alpha
ANOVA	Analysis of variance
CWs	Concrete words
CEFR	Common European framework of reference
CT	Computerised tomography
DCM	Dual coding model
DV	Dependent variable
L1	First language
χ^2	Friedman test
IV	Independent variable
IELTS	International English language testing system
IQ	Intelligence quotient
LTM	Long-term memory
MRI	Magnetic resonance imaging
U	Mann-Whitney test
M	Mean
MS	Mean square
Mdn	Median
N	Noun

η_p^2	Partial eta squared
p	Probability value
RT	Reaction time
SD	Standard deviation
n	Sample size
L2	Second language
STM	Short-term memory
SPSS	Statistical package for the social sciences
TAP	Thesis advisory panel
UKVI	United Kingdom visa
W+	Wilcoxon signed-rank test
WM	Working memory

References

Aben, B., Stapert, S., & Blokland, A. (2012). About the distinction between working memory and short-term memory. *Frontiers in Psychology*, 3, 301–314.

Altarriba, J., & Basnight-Brown, D. (2011). The representation of emotion vs. emotion-laden words in English and Spanish in the Affective Simon Task. *International Journal of Bilingualism*, 15(3), 310–328.

Altarriba J., & Bauer L. (2004). The distinctiveness of emotion concepts: A comparison between emotion, abstract, and concrete words. *The American Journal of Psychology*, 117(3), 389–410.

Altarriba, J., Bauer, L. & Benenuto, C. (1999). Concreteness, context availability, and imageability ratings and word associations for abstract, concrete and emotion words. *Behavior Research Methods, Instruments & Computers*, 31, 578–602.

Altarriba, J., & Mathis, K. M. (1997). Conceptual and lexical development in second language acquisition. *Journal of Memory and Language*, 36(4), 550–568.

Alegre, M., & Gordon, P. (1999). Frequency effects and the representational status of regular inflections. *Journal of Memory and Language*, 40(1), 41–61.

Andrews, M., Vigliocco, G., & Vinson, D. (2009). Integrating experiential and distributional data to learn semantic representations. *Psychological Review*, 116(3), 463–498.

Anooshian, L., & Hertel, P. (1994). Emotionality in free recall: Language specificity in bilingual memory. *Cognition and Emotion*, 8, 503–514.

Ashby-Davis, C. (1985). Cloze and comprehension: A qualitative analysis and critique. *Journal of Reading*, 28(7), 585–589.

Atkinson, R. & Shiffrin, R. (1968). Human memory: A proposed system and its control processes. In K. W. Spence & J. T. Spence (Eds.), *The psychology of learning and motivation,* (89–195).

Avons, S., Wright, K., & Pammer, K. (1994). The word-length effect in probed and serial recall. *The Quarterly Journal of Experimental Psychology*, 47(1), 207–231.

Ayçiçeği, A., & Harris, C. (2004). Bilinguals' recall and recognition of emotion words. *Cognition and Emotion*, 18(7), 977–987.

Baddeley, A. (2017). Short-term memory for word sequences as a function of acoustic, semantic and formal similarity. *Exploring Working Memory*, 9–14.

Baddeley, A. (2012). Working memory: Theories, models, and controversies. *Annual Review of Psychology*, 63, 1–29.

Baddeley, A. (2007). Working memory, thought and action. Oxford: Oxford University Press.

Baddeley, A. (2000a). The episodic buffer: A new component of working memory? *Trends in Cognitive Sciences*, 4, 417–423.

Baddeley, A. (2000b). *Short-term and working memory*. Oxford: Oxford handbook of memory.

Baddeley, A. (1986). *Working memory*. Oxford: Oxford University Press.

Baddeley, A, & Hitch, G. (1974). Working memory. In G. H. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory*, 47–89. NY: Academic Press.

Baddeley, A., Eysenck, M. W., & Anderson, M. C. (2009). *Memory*. East Sussex: Sussex.

Baddeley, A., Hitch, G., & Allen, R. (2019). From short-term store to multicomponent working memory: The role of the modal model. *Memory and Cognition*, 47(4), 575–588.

Baddeley, A, Thomson, N., & Buchanan, M. (1975). Word length and the structure of short-term memory. *Journal of Verbal Learning and Verbal Behavior*, 14(6), 575–589.

Balota, D., Yap, M., Cortese, M., Hutchison, K., Kessler, B., Loftis, B., & Treiman, R. (2007). The English Lexicon Project. *Behaviour Research*, 39, 445–459.

Barsalou, L. (1999). Perceptual symbol systems. *Behavioural and Brain Sciences*, 22, 577-660.

Barsalou, L., & Wiemer-Hastings, K. (2005). Situating abstract concepts. In D. Pecher and R. Zwaan (Eds.), *Grounding cognition: The role of perception and action in memory, language, and thought,* 129–163. New York: Cambridge University Press.

Basnight-Brown, D., & Altarriba, J. (2007). Differences in semantic and translation priming across languages: The role of language direction and language dominance. *Journal of Memory and Cognition*, 35, 953–965.

Bartlett, F. (1932). *Remembering: A study in experimental and social psychology*. New York: Cambridge University Press.

Bechara, A., Damasio, H., & Damasio, A. (2000). Emotion, decision making and the orbitofrontal cortex. *Cerebral Cortex*, 10(3), 295–307.

Billig, M. (1995). *Banal nationalism*. Los Angeles: Sage.

Bleasdale, F. (1987). Concreteness-dependent associative priming: Separate lexical organization for concrete and abstract words. *Journal of*

Experimental Psychology: Learning, Memory, and Cognition, 13, 582–594.

Bogaards, P. (2001). Lexical units and the learning of foreign language vocabulary. *Studies in Second Language Acquisition*, 23(3), 321–343.

Bower, G. (1970). Organizational factors in memory. *Cognitive Psychology*, 1(1), 18–46.

Bradley, M. M., & Lang, P. J. (1999). *Affective norms for English words (ANEW): Instruction manual and affective ratings*. Technical report C-1, the center for research in psychophysiology, USA, University of Florida.

Brauer, M. (1998) Stroop interference in bilinguals: The role of similarity between the two languages. In *Foreign language learning*, A. F. Healy & I. E. Bourne, (317–337).

Brysbaert, M., & Duyck, W. (2010). Is it time to leave behind the Revised Hierarchical Model of bilingual language processing after fifteen years of service? *Bilingualism: Language and Cognition*, 13(3), 359–371.

Caldwell-Harris, C., Tong, J., Lung, W., & Poo, S. (2011). Physiological reactivity to emotional phrases in Mandarin—English bilinguals. *International Journal of Bilingualism*, 15(3), 329–352.

Chan, R., Shum, D., Toulopoulou, T., & Chen, E. (2008). Assessment of executive functions: Review of instruments and identification of critical issues. *Archives of Clinical Neuropsychology*, 23(2), 201–216.

Chang, Y. (2011). The relation between time spent on the written recall task and the memory of L2 text. *Reading and Writing*, 24(8), 903–919.

Chiarello, C., Senehi, J., & Nuding, S. (1987). Semantic priming with abstract and concrete words: Differential asymmetry may be postlexical. *Brain and Languag*e, 31, 43–60.

Clark, J., & Paivio, A. (1991). Dual coding theory and education. *Educational psychology review*, 3(3), 149-210.

Clenton, J. (2015). Testing the revised hierarchical model: Evidence from word associations. *Bilingualism Language and Cognition*, 18(1), 118–125.

Clore, G., Ortony, A., & Foss, M. (1987). The psychological foundations of the affective lexicon. *Journal of Personality and Social Psychology*, 53, 751–766.

Cook, G. (1989). *Discourse in language teaching: A scheme for teacher education*. Oxford: Oxford University Press.

Cortis Mack, C., Cinel, C., Davies, N., Harding, M., & Ward, G. (2017). Serial position, output order, and list length effects for words presented on smartphones over very long intervals. *Journal of Memory and Language*, 97, 61–80.

Cowan, N. (2008). What are the differences between long-term, short-term, and working memory? *Progress in Brain Research*, 169, 323–338.

Cowan, N. Wood, N., Wood, P., Keller, T., Nugent, L., & Keller, C. (1998). Two separate verbal processing rates contributing to short-term memory span. *Journal of Experimental Psychology: General*, 127, 141–160.

Czech, D. (2016). A body of understanding: Embodiment of language in the light of current psycholinguistic and neurolinguistic research. *Orbis Linguarum*, 43, 341–357.

Daller, H., Milton, J. and Treffers-Daller, J. (2007). *Modelling and assessing vocabulary knowledge*. Cambridge; Cambridge University Press.

Davies, P. (2016). *The semantic representation of concrete and abstract words* (Unpublished doctoral thesis). University of London, London.

De Deyne, S., & Storms, G. (2015). Words associations. In by J. R. Taylor (Ed.), *The Oxford Handbook of the Word* (465–480). Oxford Handbooks in Linguistics. Oxford: Oxford University Press.

De Groot, A. (1992). Determinants of word translation. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 18, 1001–1018.

Della Rosa, P., Catricala, E., Battisti, S., Vinson, F., Vigliocco, G., & Cappa, S. (2018). How to assess abstract conceptual knowledge: Construction, standardization and validation of a new battery of semantic memory tests. *Functional Neurology*, 29(1), 47–55.

Demaree, H., Shenal, B., Everhart, D, Robinson, J. (2004). Primacy and Recency Effects Found Using Affective Word Lists. *Cognitive and Behavioral Neurology*, 17 (2), 102–108.

Dewaele, J., & Pavlenko, A. (2002). Emotion vocabulary in interlanguage. *Language Learning*, 52(2), 263–322.

Dutta, S., & Kanungo, R. (2013). *Affect and memory: A reformulation*. Pergamon Press: Oxford.

Farley, A., Ramonda, K., & Liu, X. (2012). The concreteness effect and the bilingual lexicon: The impact of visual stimuli attachment on meaning recall of abstract L2 words. *Language Teaching Research*, 16(4), 449–466.

Ferré, P. (2003). Effects of level of processing on memory for affectively valenced words. *Cognition and Emotion*, 17(6), 859–880.

Ferré, P., García, T., Fraga, I., Sánchez-Casas, R., & Molero, M. (2010). Memory for emotional words in bilinguals: Do words have the same emotional intensity in the first and in the second language? *Cognition and Emotion*, 24(5), 760–785.

Ferré, P., Sánchez-Casas, R., & Fraga, I. (2013). Memory for emotional words in the first and the second language: Effects of the encoding task. *Bilingualism: Language and Cognition*, 16(3), 495–507.

Ferré, P., Anglada-Tort, M., & Guasch, M. (2018). Processing of emotional words in bilinguals: Testing the effects of word concreteness, task type and language status. *Second Language Research*, 34(3), 371–394.

Field, A. (2013). *Discovering statistics using IBM SPSS statistics*. USA: Sage Publication.

Field, A., & Hole, G. (2002). *How to design and report experiments*. England: Sage Publication.

Field, J. (2004). *Psycholinguistics: The key concepts*. London: Psychology Press.

Fuchs, T. (2012). The phenomenology of body memory. In S. C. Koch, T. Fuchs, M. Summa, & C. Mueller (Eds.), *Body memory, metaphor and movement* (9–22). Amsterdam: John Benjamins Publishing Company.

Garrido, M & Prada, M. (2018). Comparing the valence, emotionality and subjective familiarity of words in a first and a second language. *International Journal of Bilingual Education and Bilingualism*, 1–17

Glenberg, A. & Robertson, A. (1999). Indexical understanding of instructions. *Discourse Processes*, 28(1), 1–26.

Gorman, A. (1961). Recognition memory for nouns as a function of abstractness and frequency. *Journal of Experimental Psychology*, 61(1), 23–29.

Griffiths, T, Steyvers, M., & Tenenbaum, J. (2007). Topics in semantic representation. *Psychological Review*, 114(2), 211–244.

Guthrie, W. (2012). *The Greek Philosophers*: From Thales to Aristotle. London: Routledge.

Hampton, J., & Moss, H. (2003). Concepts and meaning: Introduction to the special issue on conceptual representation. *Language and Cognitive Processes*, 18, 505–512.

Harris, C., Gleason, J., & Ayçiçeği, A. (2006). *When is a first language more emotional? Psychophysiological evidence from bilingual speakers*. Turkey: Istanbul Publisher Company.

Hedblom, M., Kutz, O., & Neuhaus, F. (2016). Image Schemas as Families of Theories. Cognitive Systems Research, 39, 42–57.

Hulme, C., Stuart, G., Brown, G. & Morin, C. (2003). High-and low-frequency words are recalled equally well in alternating lists: Evidence for associative effects in serial recall. *Journal of Memory and Language*, 49(4), 500–518.

James, W. (1890). *The principles of psychology*. New York: Henry Holt Company.

Johnson-Laird, P. (1983). *Mental models: Towards a cognitive science of language, inference and consciousness*. Cambridge: Cambridge University Press.

Johnson, M. (1987). *The body in the Mind. The bodily basis of meaning, imagination, and reasoning*. Chicago: The University of Chicago Press.

Kensinger, E. (2008). Age Differences in Memory for Arousing and Non-arousing Emotional Words. *The Journals of Gerontology. Psychological Sciences and Social Sciences*, 63(1), 13–18.

Kiss, G. (1968). Words, associations and networks. *Journal of Verbal Learning and Verbal Behaviours*, 7, 707–713.

Khokhlova, N. (2014). Understanding of abstract nouns in linguistic disciplines. *Procedia of Social and Behavioural Sciences*, 136, 8–11.

Kline, P. (2000). *Handbook of psychological testing*. London: Routledge.

Kousta, S., Vinson, D., & Vigliocco, G. (2009). Emotion words, regardless of polarity, have a processing advantage over neutral words. *Cognition*, 112, 473–481.

Kousta, S., Vigliocco, G., Vinson, D., Andrews, M., & Del Campo, E. (2011). The representation of abstract words: Why emotion matters. *Journal of Experimental Psychology: General*, 140, 14–34.

Kroll, J., & Potter, M. (1984). Recognizing words, pictures, and concepts: A comparison of lexical, object, and reality decisions. *Journal of Verbal Learning and Verbal Behavior*, 23, 39–66.

Kroll, J., & Stewart, E. (1994). Category interference in translation and picture naming: Evidence for asymmetric connections between bilingual memory representations. *Journal of Memory and Language*, 33, 149–174.

Kroll, J., Michael, E., & Tokowicz, N. (2002). The development of lexical fluency in a second language. *Journal of Second Language Research*, 18(2), 137–171.

Lakoff, G. (1987). *Women, fire, and dangerous things*. Chicago: University of Chicago.

Lakoff, G., & Johnson, M. (1980a). *Metaphor we live by*. Chicago: University of Chicago Press.

Lakoff, G., & Johnson, M. (1980b). The metaphorical structure of the human conceptual system. *Cognitive Science*, 4(2), 195–208.

Larson-Hall, J. (2015). *A Guide to Doing Statistics in Second Language Research Using SPSS and R*. New York: Routledge.

Langdridge, D., & Hagger-Johnson, G. (2009). *Introduction to research methods and data analysis in psychology*. London: Pearson Education.

Laufer, B. (1997). What's in a word that makes it hard or easy? Some intralexical factors that affect the learning of words. In *Vocabulary: Description, acquisition and pedagogy* (140–155). Cambridge: Cambridge University Press.

Lockhart, R. (2000). *Methods of memory research*. Oxford handbook of memory, 45–57.

Maalej, Z. (1990). *Metaphor in political and economic texts*. (Unpublished PhD thesis). University of Manouba:Tunisia.

Majerus, S., & Van der Linden, M. (2003). Long-term memory effects on verbal short-term memory: A replication study. *British Journal of Developmental Psychology*, 21(2), 303–310.

Marmolejo-Ramos, F., & Dunn, J. (2013). On the activation of sensorimotor systems during the processing of emotionally laden stimuli. *Universitas Psychologica*, 12(5), 1515–1546.

Martin, K., & Ellis, N. (2012). The roles of phonological short-term memory and working memory in L2 grammar and vocabulary learning. *Studies in Second Language Acquisition*, 34, 379–413.

Mazzuca, C., Barca, L., & Borghi, A. (2017). The peculiarity of emotional words: a grounded approach. *Rivista internazionale di Filosofia e Psicologia*, 8(2), 124–133.

Miller, G. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, 63, 81–97.

Milton, J. (2009) *Measuring Second Language Vocabulary Acquisition*. Bristol: Multilingual Matters.

Morton, J. (1969). Interaction of information in word recognition. *Psychological Review*, 76, 165–178.

Naccache, L., Adam, C., Baulac, M., Clemenceau, S., Cohen, L., Dehaene, S., Gaillard, R. & Hasboun, D. (2005). A direct intracranial record of emotions evoked by subliminal words. *Proceedings of the National Academy of Sciences*. 102, 7713–7717.

Nairne, J. (2002). Remembering over the short-term: The case against the standard model. *Annual Review of Psychology*, 53, 53–8.

Nation, I. (2001) *Learning vocabulary in another language*. Cambridge; Cambridge University Press.

Nation, I. (2013a). *Teaching & learning vocabulary*. Boston: Heinle Cengage Learning.

Nation, I. (2013b). *Learning vocabulary in another language*. New York: Cambridge University Press.

Nelson, D., & Schreiber, T. (1992). Word concreteness and word structure as independent determinants of recall. *Journal of Memory and Language*, 31, 237–260.

Norris, D. (2017). Short-term memory and long-term memory are still different. *Psychological Bulletin*, 143(9), 992–1009.

Novitskaya I. (2006). Philosophic and linguistic interpretation of abstract nouns. *Vestnik*, 4(55), 7–11.

Oakley, T. (2007). Image schemas. *Oxford handbook of cognitive linguistics*, 214–235.

Okano, H & Balaban, E. (2000). Learning and memory. *The Second Annual Japanese- American Frontiers of Science Symposium*, 7, 23–97.

Oswald, F., McAbee, S., Redick, T., & Hambrick, D. (2015). The development of a short domain-general measure of working memory capacity. *Behaviour Research Methods*, 47(4), 1343–1355.

Paivio, A., Rogers, T., and Smythe, P. (1968). Why are pictures easier to recall than words? *Psychonomic Science*, 11(4), 137–138.

Paivio, A., & Csapo, K. (1969). Concrete image and verbal memory codes. *Journal of Experimental Psychology*, 80, 279–285.

Paivio, A. (1971a). *Imagery and verbal processes*. New York: Holt, Rinehart, and Winston.

Paivio, A. (1971b). Short-term sequential memory for pictures and words. *Psychonomic Science*, 24(2), 50–51.

Paivio, A. (1973). Picture superiority in free recall imagery or dual coding. *Journal of Cognitive Psychology*, 5, 176–206.

Paivio, A. & Desrochers, A. (1980). A dual-coding approach to bilingual memory. *Canadian Journal of Psychology*, 34, 388–99.

Paivio, A., & Lambert, W. (1981). Dual coding and bilingual memory. *Journal of Verbal Learning and Verbal Behavior*, 20, 532–539.

Paivio, A. (1991). *Images in mind: The evolution of a theory*. UK: Harvester Wheatsheaf.

Paivio, A., (2006). 'Dual coding theory and education'. In Draft paper for the conference on *Pathways to literacy achievement for high poverty children*, The University of Michigan School of Education, USA (Sept 2006).

Pavlenko, A. (2005). *Emotions and multilingualism.* Cambridge: Cambridge University Press.

Pavlenko, A. (2008a). Emotion and emotion-laden words in the bilingual lexicon. *Bilingualism: Language and Cognition*, 11, 147–64.

Pavlenko, A. (2008b). Structural and conceptual equivalence in the acquisition and use of emotion words in a second language. *The Mental Lexicon*, 3(1), 92–121.

Peterson, E., Deary, I., & Austin, E. (2003). The reliability of Riding's cognitive style analysis test. *Personality and Individual Differences*, 34(5), 881–891.

Poirier, M., & Saint-Aubin, J. (1995). Memory for related and unrelated words: Further evidence on the influence of semantic factors in immediate serial recall. *The Quarterly Journal of Experimental Psychology*, 48(2), 384–404.

Pufahl, A., & Samuel, A. G. (2014). How lexical is the lexicon? Evidence for integrated auditory memory representations. *Cognitive Psychology*, 70, 1–30.

Pulvermüller., F. (2013). How neurons make meaning: Brain mechanisms for embodied and abstract-symbolic semantics. *Trends in Cognitive Sciences*, 13(9), 458–470.

Quinlan P., & Dyson., B (2008). *Cognitive psychology*. England: Pearson Publisher.

Radvansky, G. (2006) *Human memory*. England: Pearson Education Group.

Rhee, S. (2002). Semantic changes of English preposition against: A grammaticalization perspective. *Linguistic Research*, 38, 563–583.

Richardson, J. (1996). Evolving issues in working memory. *Working Memory and Human Cognition*, 120–154.

Rogers, V., Playfoot, D. & Milton, J. (2018). The Second Language Lexicon. In C. Wright, T. Piske & M. Young-Scholten (Ed.), *Mind Matters in SLA* (117-136). Bristol, Blue Ridge Summit: Multilingual Matters.

Rosch, E., & Lloyd, B. (1978). *Cognition and categorization*. New Jersey: Lawrence Erlbaum Associates.

Schwanenflugel, P., & Shoben, E. (1983). Differential context effects in the comprehension of abstract and concrete verbal materials. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 9(1), 82–102.

Schwanenflugel, P., Harnishfeger, K., & Stowe, R. (1988). Context availability and lexical decision for abstract and concrete words. *Journal of Memory and Language*, 27, 499–520.

Schwanenflugel, P., & Stowe, R. (1989). Context availability and the processing of abstract and concrete words in sentences. *Reading Research Quarterly*, 24, 114–126.

Schwanenflugel, P. (1991). Why are abstract concepts hard to understand? In P. J. Schwanenflugel (Ed.), *The psychology of word meanings*, 223–250.

Schwanenflugel, P. (1992). Context availability and the recall of abstract and concrete words. *Memory and Cognition*, 20(1), 96–104.

Scorolli, C., Binkofski, F., Buccino, G., Nicoletti, R., Riggio, L., & Borghi, A. (2011). Abstract and concrete sentences, embodiment, and languages. *Frontiers in Psychology*, 2, 227–237.

Scorolli, C., Jacquet, P., Binkofski, F., Nicoletti, R., Tessari, A., & Borghi, A. (2012). Abstract and concrete phrases processing differentially modulates cortico-spinal excitability. *Brain Research*, 1488, 60–71.

Shaver, P., Schwartz, J., Kirson, D., & O'Connor, C. (1987). Emotion knowledge: Further exploration of a prototype approach. *Journal of Personality & Social Psychology*, 52, 1061–1086.

Siakaluk, P., Pexman, P., Sears, C., Wilson, K., Locheed, K., & Owen, W. (2008). The benefits of sensorimotor knowledge: Body–object interaction facilitates semantic processing. *Cognitive Science*, 32, 591–605.

Siakaluk, P., Knol, N., & Pexman, P. (2014). Effects of emotional experience for abstract words in the Stroop task. *Cognitive Science*, 38(8), 1698–1717.

Siakaluk, P., Newcombe, P., Duffels, B., Li, E., Sidhu, D., Yap, M., & Pexman, P. (2016). Effects of emotional experience in lexical decision. *Frontiers in Psychology*, 7, 1157–1177.

Squire, L. (1987). *Memory and brain*. New York: Oxford University Press.

Sternberg, J. (2006). *Cognitive psychology*. USA: Thomson Wadsworth.

Stevenson, A. (2010). *Oxford Dictionary of English*. Oxford University Press.

Toglia, M., & Battig, W. (1978). *Handbook of semantic word norms*. New Jersey: Lawrence Erlbaum Associates.

Tracy, D. (1979). *Metaphor and religion*. Chicago: The University of Chicago Press.

Tulving, E. (1972). *Episodic and semantic memory*. Oxford: Academic Press.

Tulving, E. (2000). *Concepts of memory*. Oxford: The Oxford handbook of memory.

Vigliocco, G., Meteyard, L., Andrews, M., & Kousta, S. (2009). Toward a theory of semantic representation. *Language and Cognition*, 1(2), 219–247.

Vigliocco, G., Della Rosa, P., Vinson, D. P., Kousta, S., Devlin, J., & Cappa, S. (2010). The neural representation of abstract words. *Cerebral Cortex,* 24(7), 1767–1777.

Vigliocco, G., Kousta, S., Vinson, D., Andrews, M., & Del Campo, E. (2013). The representation of abstract words: What matters? Reply to Paivio's (2013) comment on Kousta et al. (2011). *Journal of Experimental Psychology*, 142(1), 288–291.

Wegner, D. (1987). Transactive memory: A contemporary analysis of the group mind. In *Theories of group behavior* (185–208). New York: Springer.

Wiemer-Hastings, K. (1998). Abstract noun classification: using a neural network to match word context and word meaning. *Behavior Research Methods, Instruments, & Computers*, 30(2), 264–271.

Wiemer-Hastings, K., Krug, J., & Xu, X. (2001). Imagery, context availability, contextual constraint and abstractness. *Proceedings of the Annual Meeting of the Cognitive Science Society*, 23, (23), 1–7.

Wilks, C., & Meara, P. (2007). Implementing graph theory approaches to the exploration of density and structure in L1 and L2 word association networks. In *Modelling and assessing vocabulary knowledge* (167–181). Cambridge University Press.

Wlotko, E., & Federmeier, K. (2007). Finding the right word: Hemispheric asymmetries in the use of sentence context information. *Neuropsychologia*, 45(13), 3001–3014.

Yoo, J., & Kaushanskaya, M. (2016). Serial-position effects on a free-recall task in bilinguals. *Memory*, 24(3), 409–422.

Appendix A

Consent Form of Data Collection

Department of Language and Linguistics Science
Heslington, York, YO10 5DD, UK
Linguistics@york.ac.uk

The Role of Emotionality in the Organisation of English Abstract Words on Second Language Lexicon

***Lead researcher**: Yasir Almukhaizeem*

Consent form (English version)

This form asks you to state whether you agree to take part in the study. Please read and answer every question. If there is anything you do not understand or if you want more information, please ask the researcher.

Have you read and understood the information leaflet about the study?	Yes❐ No ❐
Have you had an opportunity to ask questions about the study and have these been answered satisfactorily?	Yes❐ No ❐
Do you understand that the information you provide will be held in confidence by the research team, and your name or identifying information about you will not be mentioned in any publication?	Yes❐ No ❐

Do you understand that you may withdraw from the study at any time before the end of the data collection session without giving any reason and that, in such case, all your data will be destroyed?	Yes❒ No ❒
Do you understand that the information you provide may be kept after the completion of the current project to be used in future research on language?	Yes❒ No ❒
Do you agree to take part in the study?	Yes❒ No ❒
If yes, do you agree to your interview being recorded on video?	Yes❒ No ❒
Do you agree to researchers using the excerpts from your audio recordings in presentations or in teaching without disclosing your real name? (You may take part in the study without agreeing to this).	Yes❒ No ❒
Do you agree to allow the researcher to keep your contact details after the end of the current project in order that s/he may contact you in the future about possible participation in other studies? (You may take part in the study without agreeing to this).	Yes❒ No ❒

Your Name (in BLOCK letters): ____

Your signature: ____

Researcher's name: ***Yasir Almukahaizeem (email: ysaa500@york.ac.uk)***

Date: ____

تأثير مستوى العاطفة على تمثيل الكلمات المجردة في المعجم ثنائي اللغة

(النسخة العربية)

استمارة الموافقة على اجراء دراسة

أخي المشارك، أنشئ هذا النموذج من أجلك لكي توضح مدى موافقتك او رفضك للمشاركة في الدراسة. الرجاء التكرم بقراءة كل الأسئلة والاجابة عليها. إذا واجهتك أي غموض او اردت الاستزادة من المعلومات، من فضلك التوجه بالسؤال للباحث.

هل قرأت وفهمت نشرة المعلومات الموضحة للدراسة؟ نعم❒ لا❒

هل أتيحت لك الفرصة لطرح أسئلة عن الدراسة، وهل تمت الإجابة عليها بشكلٍ مرضٍ؟ نعم❒ لا❒

هل أنت على دراية بأن المعلومات التي ستقدمها لنا ستحفظ بشكل متقن من قبل فريق البحث، ولن يتم ذكر اسمك او المعلومات الخاصة بك في أي وعاء نشر ؟ نعم❒ لا❒

هل تعلم بأنه بإمكانك الانسحاب من المشاركة في هذه الدراسة في أي وقت قبل نهاية جمع البيانات دون إبداء أي سبب، وأنه في مثل هذه الحالة سيتم التخلص من جميع البيانات الخاصة بك؟ نعم❒ لا❒

هل تعلم ان المعلومات التي ستقدم لنا قد تبقى مدة المشروع او بعده لاستخدامها في البحوث المستقبلية في اللغة؟ نعم❒ لا❒

هل موافق على المشاركة في الدراسة؟ نعم❒ لا❒

إذا كانت الاجابة بنعم، هل توافق على تسجيل مقابلتك صوتياً؟ نعم❒ لا❒

هل توافق على أخذ مقتطفات من تسجيلك الصوتي وذلك لاستخدامها من قبل الباحث في العروض التقديمية أو التدريس، وبدون الكشف عن اسمك. نعم❒ لا❒

(بإمكانك المشاركة في الدراسة بدون هذه الموافقة)

هل توافق على الاحتفاظ ببيانات الاتصال بك بعد نهاية هذا المشروع؛ وذلك حتى يتمكن التواصل معك في المستقبل بغرض المشاركة المتحملة منك في دراسات أخرى نعم❐ لا❐

(بإمكانك المشاركة في الدراسة بدون هذه الموافقة)

الاسم: ____________________

التوقيع: ____________________

التاريخ ____________________

اسم الباحث: **ياسر بن سعد المخيزيم**

Announcement

(Recruiting Monolingual Participants)

My name is Yasir and I am a PhD candidate at the University of York, York.

I am conducting research for my dissertation in linguistics.

What do I have to do if I take part?

- Attend one-hour session which involved remembering and rating words
- You will be asked to wear a headset.

Participants are required to be

- Male native speakers of English
- Speak only English
- Age (20-30)
- Be healthy with no brain lesions
- Be a right-hander.

Participation is voluntary and you will be reimbursed 10 pounds.

For more information about this study, please contact me via:

Ysaa500@york.ac.uk

Announcement

(Recruiting L2 Participants)

My name is Yasir and I am a PhD candidate at the University of York, York.

I am conducting research for my dissertation in linguistics.

What do I have to do if I take part?

- Attend one-hour session which involved remembering and rating words
- You will be asked to wear a headset.

Participants are required to be

- Age (20-30)
- Be healthy with no brain lesions
- Be a right-hander
- Beginners or advanced of English
- Have recent IELTS UKVI.

Participation is voluntary and you will be reimbursed 10 pounds.

For more information about this study, please contact me via:

Ysaa500@york.ac.uk

Appendix B

Long-term Memory Task (1)

Participant: **( )** **Group:** **( )**

Language proficiency level for L2 group Advanced ( ) Beginner ( )

Prerequisite steps

- Consent form ( )
- Information sheet ( )
- Eligibility of the participant (e.g., *age, gender, right-handers*) ( )

Instructions

- You will be given a list of sentences; your job is to complete the sentences with the appropriate word to fill in the blank. Each sentence should be meaningful and easy to understand;
- Write only **ONE WORD** which completes each sentence correctly;
- Do not use words in this experiment to fill in the blank(s); and
- If you have a question, please do not hesitate to ask the researcher directly.

Complete the following sentences with an appropriate word. Sentences must be clear and meaningful.

1. People were warned to be on..………….**(emotional word)**[12] for possible suspicious threats in Mexico starting from tonight.

[12] The type of AWs is written between brackets in each sentence. They were added here for further clarification. During the actual task, these brackets were omitted.

2. Earthquake, in addition to the volcano, is a natural……….…**(emotional word)** that affects our planet. It has severed effect as heavy rains and volcanoes.

3. In Christianity, Islam, and Judaism, ..……… **(emotional word)** is the most powerful evil spirit as opposed to an angel.

4. Gem was crying in the hospital last night. She believed it was the doctor's ………….**(emotional word)** that Peter died.

5. Paul has a lovely………… **(non-emotional word)** in dealing with people. He always shows respect and appreciation even if he does not like the person.

6. Surely, it is wrong to try to impose Western…………..**(non-emotional word)** on people. Every country is different values, and we need to understand that.

7. For many television viewers, the dividing line between fact and…………..**(non-emotional word)** is becoming increasingly difficult to distinguish. Some TV shows have a lot of lies.

8. Fish, birds, insects, reptiles, mammals, cats, and snakes are all creatures. They are categorized as…………... (**CW**).

9. I was talking to my boss in the location. We expected the main……….…..(**CW**) to be finished ahead of schedule. Other facilities can be finished later.

10. According to the recipe, you should bake …………(**CW**) in a hot oven, at about 220°C, for 30 minutes and then cover it with cream and nuts.

11. It's sad that he never gets any real…………(**emotional word**) out of his vacation. I think he should have pleasure, and not answering calls or emails in his fun time.

12. I was sad after the party last night. I could not understand why no one ever laughed at my…………… (**emotional word**).

13. For most citizens,…………….(**emotional word**) means the freedom to practice their own choices and religious beliefs with no segregation.

14. To my………….. (**non-emotional word**), there are only two Indian restaurants that serve good food in this town. You should go there sometime.

15. Come on, James! Of course, I'm telling the truth. Why do you have such a suspicious …………… (**non-emotional word**)?

16. You are doing well at this level! You have achieved a complete ……………. (**non-emotional word**) of math skills so far, and I anticipate more progress in the future.

17. Both performances were excellent! This is simply my……………(**non-emotional word**) when I saw both of them for the first time.

18. The police officer told us that the diamond thief ran away with his partners and gave the assistant buyer worthless fake…………...(**CW**).

19. Look at her! She clapped her hand over and covered her………….(**CW**) to try to stop herself from laughing. She must be very happy.

20. James had a medical problem on his face. He had a recent plastic surgery on his…………..(**CW**) to straighten it a bit.

21. Coming to your office on time as well as doing your responsibilities efficiently grantees that you will get a ………… (**emotional word**) soon.

22. Her mother felt a great sense of………….. (**emotional word**) as she watched her son invited to get the award for his success in business.

23. Honestly speaking, I've never witnessed such two different extremes of wealth and………….. (**emotional word**) in my life.

24. Excuse me sir! I would like to place an…………..(non-emotional word) for a large table for my birthday party next week.

25. Television has an important……………..(**non-emotional word**) to play in popularizing new ideas and projects in the society.

26. The machine's…………(**non-emotional word**) is recorded, and its results will be sent to James including performance, maintenance, and potential costs.

27. Newspaper and………………(**CW**) have various types of paper-based advertisements either with/without pictures.

28. I must admire Shell's daily routine. She drives her children to…………(**CW**) every morning and go shopping.

29. Good morning, officer! I need your assistance in parking my car. I can't go down to that…………(**CW**) because there is a "No Entry" sign.

30. In addition to fat, people generally eat too much………….(**CW**) in their food. In fact, they do not know it is the main cause of high blood pressure.

Appendix C

Long-term Memory Task (2)

الرقم ( )

تعليمات الاختبار

أجب عن الأسئلة التالية كما هو موضح. الوقت المسوح لإتمام الإجابة على كل الأسئلة هو عشرون دقيقة.

اختر إجابة واحدة فقط من الخيارات التالية. يجب ان يمثل الاختيار المعنى المقصود للكلمة.

1. فعل يستخدم للتشجيع على حدوث شيء ما أو تطويره.

- Enjoyment
- Joke
- Promotion

2. شيء ما يمثل قصة مضحكة أو حيلة يقال أو يُفعل لإضحاك الناس.

- Joke
- Pride
- Knowledge

3. حرية العيش كما يحلو لك والذهاب إلى حيث تريد.

- Fault
- Character
- Liberty

4. الشعور بالسعادة أو السرور.

- Pride
- Joke
- Enjoyment

5. الشعور بالسعادة والرضا الذي تحصل عليه لأنك أنت أو الأشخاص المرتبطين بك قد فعلت أو حصلت على شيء جيد.

- Pride
- Joke
- Enjoyment

6. تحذير للناس بالاستعداد للتعامل مع شيء خطير.

- Disaster
- Alert
- Mind

7. الحدث الذي ينتج عنه ضرر جسيم أو ضرر أو وفاة أو صعوبة جسيمة.

- Disaster
- Poverty
- Fault

8. كائن شرير غالبًا ما يتم تمثيله في شكل بشري ولكن بذيل وقرون.

- Devil
- Poverty
- Disaster

9. خطأ وخاصة الشيء الذي تتحمل المسؤولية عنه.

- Disaster
- Fiction
- Fault

10. حالة الفقر المدقع.

- Poverty
- Pride
- Mastery

11. شيء يعيش ويتحرك ولكن ليس بشراً.

- Animal
- Mouth
- Disaster

12. هيكل بجدران وسقف مثل منزل أو مصنع.

- Opinion
- Building
- Order

13. طعام حلو مصنوع من خليط من الطحين والبيض والدهون والسكر.

- Cake
- Mouth
- Salt

14. حجر كريم يستخدم لتزيين الأشياء الثمينة.

- Jewelry
- Pride
- Mastery

15. فتحة في وجه إنسان أو حيوان، تتكون من الشفتين والمسافة بينهما، أو الفراغ الموجود خلف الأسنان واللسان.

- Nose
- Mouth
- Pride

16. جزء الوجه الذي يبرز فوق الفم والذي من خلاله تتنفس وتشم.

- Nose
- Mouth
- Pride

17. نوع من الكتب الرقيقة ذات الصفحات الكبيرة والغلاف الورقي الذي يحتوي على مقالات وصور ويتم نشره كل أسبوع أو شهر.

- Character
- Opinion
- Magazine

18. فترة حياتك التي تذهب فيها إلى المدرسة او أنشطة التدريس والتعلم التي تحدث في المدرسة.

- School
- Building
- Enjoyment

19. طريق في مدينة أو بلدة بها مبان عادة ما تكون قريبة من بعضها على طول أحد الجانبين أو كلاهما.

- Building
- Street
- Mastery

20- مادة بيضاء شائعة توجد في مياه البحر وفي باطن الأرض، تستخدم بشكل خاص لإضفاء نكهة على الطعام أو لحفظه.

- Salt
- Cake
- Order

ترجم الكلمات التالية إلى اللغة العربية وذلك بكتابة كلمة واحدة فقط تترجم المعنى.

1. Character...
2. Culture…………...
3. Fiction…………...
4. Knowledge…………...
5. Mind…………..
6. Mastery…………..
7. Opinion…………..
8. Order …..
9. Position…………..
10. Response…………..